PCT/GB 9 9 / 0 0 8 4 9

The Patent Office

INVESTOR IN PEOPLE

EA3U

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP9 1RH

REC'D 2 8 APR 1999
WIPO PCT

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

I also certify that by virtue of an assignment registered under the Patents Act 1977, the application is now proceeding in the name as substituted.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

Signed Andrews.

Dated 09 APR 1999

The Patent Office

9806093.2.

By virtue of a direction given under Section 30 of the Patents Act 1977, the application is proceeding in the name of

St Georges Hospital Medical School
Crammer Terrace
London
SW17 0RE

Patents Form 1/77

Patents Act 1977
(Rule 16)

The Patent Office

24MAR98 E347648-1 D00192
P01/7700 25.00 - 9806093.2

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form))

The Patent Office

Cardiff Road
Newport
Gwent NP9 1RH

1. Your reference — P.74368 DMG/IJB

2. Patent application number *(The Patent Office will fill this in)* — 9806093.2 — 20 MAR 1998

3. Full name, address and postcode of the or of Each applicant *(underline all surnames)*

(1) JOHN HERMON-TAYLOR
(2) TIMOTHY JOHN BULL
DEPARTMENT OF SURGERY
ST. GEORGES HOSPITAL MEDICAL SCHOOL
CRANMER TERRACE
LONDON SW17 0RE
UNITED KINGDOM

Patents ADP number *(if you know it)*
(1) 6988657001
(2) 7044510001

If the applicant is a corporate body, give the country/state of its incorporation — country/state of incorporation

4. Title of the invention — "DIAGNOSTICS AND VACCINES FOR ~~MYCOBACTERIUM~~ MYCOBACTERIAL PARATUBERCULOSIS INFECTIONS OF ANIMALS AND HUMANS"

5. Name of your agent *(if you have one)* — J A KEMP & CO

"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*

14 SOUTH SQUARE
GRAY'S INN
LONDON WC1R 5LX

Patents ADP number *(if you know it)* — 26001

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number

Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application	Date of filing *(day / month / year)*

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer "Yes" if:*
   a) any applicant named in part 3 is not an inventor, or
   b) there is an inventor who is not named as an applicant, or
   c) any named applicant is a corporate body:
   See note (d))

Patents Form 1/77

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form	
Description	55
Claim(s)	7
Abstract	
Drawing(s)	

10. If you are also filing any of the following, state how many against each item.

Priority documents	
Translations of priority documents	
Statement of inventorship and right to grant of a patent *(Patents Form 7/77)*	
Request for preliminary examination and search *(Patents Form 9/77)*	
Request for substantive examination *(Patents Form 10/77)*	
Any other documents *(please specify)*	

11. I/We request the grant of a patent on the basis of this application

Signature J M Goldin. Date 20 MARCH 1998

12. Name and daytime telephone number of person to contact in the United Kingdom

D M GOLDIN
0171 405 3292

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.

b) Write your answers in capital letters using black ink or you may type them.

c) If there is not enough space for all the relevant details on any part of this form, please continue of a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.

d) If you have answered "Yes" Patents Form 7/77 will need to be filed.

e) Once you have filled in the form you must remember to sign and date it.

f) For details of the fee and ways to pay please contact the Patent Office.

DIAGNOSTICS AND VACCINES FOR MYCOBACTERIAL INFECTIONS OF ANIMALS AND HUMANS.

This invention relates to the protein, *Mycobacterium paratuberculosis* acylase (*mpa*) and the gene encoding *mpa*, which we have identified in the pathogen *Mycobacterium paratuberculosis* (also designated *Mycobacterium avium* subspecies *paratuberculosis*) *Mptb*, and to their use in the diagnosis of *Mptb* infections in animals and humans, as well as their use as components of vaccines for the prevention and treatment of diseases caused by *Mptb*. This invention also relates to the continuous uninterrupted *mpa* gene as we have found it to occur in some other pathogenic isolates of *M.avium*. We recognise the importance of an intact *mpa* gene as a determinant of pathogenicity in *Mptb* and other pathogenic *M.avium* species. Thus the invention also provides attenuated strains of normally pathogenic mycobacteria in which *mpa* has been inactivated, for use as vaccines.

Mptb is a pathogenic mycobacterium and a member of the group of mycobacteria called *M.avium-intracellulare* (MAIC). *Mptb* causes Johne's disease, a chronic inflammation of the intestine of a broad range of different types in many species of animals including primates. *Mptb* also causes Crohn's disease in humans and other chronic inflammatory diseases such as sarcoidosis. Johne's disease is widespread in Europe and North America as well as elsewhere, and is a major problem in both domestic and wild animals causing substantial economic losses. Crohn's disease in humans is increasing in frequency in Western Europe, North America and elsewhere, and is a major cost to health services.

The diagnosis, prevention and treatment of *Mptb* infections present major problems for veterinary and human medicine, as well as for public health. The development of effective

new diagnostics and vaccines for the recognition, prevention and/or treatment of *Mptb* disease, depend upon identifying specific genes and their products within *Mptb* that are highly specific to this organism and associated with pathogenicity. *Mptb* is very similar to other non-pathogenic members of the MAIC, but some of its genes such as IS*900* and the genes of the GS element, are known to be specific or highly selective for this organism and are associated with pathogenicity. Thus, IS*900* and GS element polynucleotides and polypeptides are useful in diagnosis and as components of vaccines against *Mptb*.

We have now identified a new gene which occurs naturally within the genome of *Mptb* and which encodes a new protein, *mpa*. This is termed the *mpa* gene of the invention. The nucleotide sequence of the *mpa* gene is given in SEQ ID No. 1 and the amino acid sequence of the *mpa* polypeptide it encodes is given in SEQ ID No. 2. *Mpa* is specific for *Mptb* and some other pathogenic *MAIC*. Unlike the genes in the GS element, homologues of the *mpa* gene are not present in pathogenic *Mycobacterium tuberculosis*. Furthermore, we have found that the equivalent gene to *mpa* in the closely related pathogenic *M.avium* subsp. *silvaticum (Mavs)* is knocked out by the presence of an IS*21*- like insertion sequence, IS*1612*. The nucleotide sequence of IS*1612* is given in SEQ ID No.3 and its complement in SEQ ID No. 4. Nucleotides 1856-2543 of SEQ ID No. 3 and nucleotides 1-688 of SEQ ID No. 4 are new and constitute further aspects of the invention. All or part of the polynucleotide sequence of IS*1612* may be used to deactivate *mpa* in *Mptb*.

We have found that the *mpa* gene encodes a polypeptide (SEQ ID No.2) whose function is to acylate cell wall components, particularly sugars, in *Mptb*. More precisely *mpa* functions in conjunction with the five polypeptides encoded by the ORFs present in the GS region of *Mptb*. The five polypeptides encoded by the GS are: *gsa* a sugar transferase, *gsb*A and *gsb*B

which function in tandem to produce fucose, *gsc* a methylase, and *gsd* a fucosyl transferase. These five polypeptides of GS serve to provide glycosyl peptidolipids (GPL) which comprise a fucosyl moiety and which are located in the mycobacterial cell wall. The methylase *gsc* acts to methylate, among other groups, the fucosyl moiety of the GPL making it less recognisable by the host's immune system. The acylase, *mpa*, acts as an acetylating agent on, among other groups, the fucosyl moiety. Acetylation of this GPL broadens the range of animals and host cells which can be infected by *Mptb*.

Acetylation of surface polysaccharides is important in host cell recognition. For example, it is known that de-acetylation of the terminal fucose of GPL's from *M.avium* MAC serotype 9 abolishes rabbit anti-serotype 9 antibody agglutination indicating that acetylation plays a key role in strain variability amongst bacteria and may be directly attributable as a virulence factor. This is not the case for *Mycobacterium avium* subsp. *avium* serotype 2 GPL. Acetylation of rhamnose in GPL's of *Mycobacterium smegmatis* induces resistance to mycobacteriophage D4 by inducing conformational changes that destroy the phage attachment site. In *Mptb*, the *mpa* gene is responsible for modification of terminal sugar residues in *Mptb* GPL's which are critically important in determining cell surface recognition and receptor binding and are important determinants of pathogenicity.

Acetylation of GPL's is also known to be critical in destroying antibody epitopes in *S.typhimurium* (Slauch J.M., Lee A.A., Mahan M.J., Mekalanos J.J., J.Bacteriol. 1996, 178, 5904-5909).

Furthermore, acetylation of GPL's is known to increase the virulence of bacteria such as *S.fexneri* (Clark C.A., Beltrame J. and Manning P.A., Gene 1991, 107, 43-52).

Due to its specificity to *Mptb* and other pathogenic *M.avium* species and the involvement of *mpa* in the biosynthesis and modification of cell wall components, *mpa* polypeptide can be used as a target for immunodiagnostic tests for the recognition of *Mptb* infections in animals and humans. Furthermore, immunisation of animals or humans with *mpa* derived products such as naked *mpa* DNA in suitable expression constructs, or *mpa* polypeptide or fragments thereof with adjuvants and carriers, will enhance the immunological resistance of animals and humans to *Mptb* infections.

It is also an objective of the present invention to provide new methods for diagnosing Johne's disease in animals and Crohn's disease in humans using immunoassays based on antibody and cell mediated immune reactivity to *mpa*. It is a further objective of the present invention to prevent and/or treat *Mptb* infections in animals and humans by vaccination using the polynucleotides or polypeptides of the *mpa* gene either alone or together with other polynucleotides and polypeptides associated with pathogenicity, such as those of GS and IS*900* known in the art. We recognise that vaccination using naked *mpa* DNA together with naked GS DNA both in suitable expression constructs will result in the provision of GPL structures on the cell surface of recipient cells, which GPL will resemble those of *Mptb* resulting in the generation of protective immunity.

Accordingly, the present invention provides a polynucleotide encoding a protein that comprises mycobacterium paratuberculosis acylase (*mpa*), or a fragment or homologue thereof having *mpa* activity.

The invention also provides a polynucleotide selected from:

(a) a polynucleotide comprising the nucleotide sequence set out in SEQ ID No. 1 or the complement thereof;

(b) a polynucleotide comprising a nucleotide sequence capable of hybridising to a fragment of the nucleotide sequence set out in SEQ ID No. 1, the fragment having the nucleotide sequence of nucleotides 210-1335 of SEQ ID No. 1;

(c) a polynucleotide comprising a nucleotide sequence capable of hybridising to the complement of a fragment of the nucleotide sequence set out in SEQ ID No. 1, the fragment having the nucleotide sequence of nucleotides 210-1335 of SEQ ID No. 1;

(d) a polynucleotide comprising a polynucleotide sequence which is degenerate as a result of the genetic code to the polynucleotide of SEQ ID No. 1 or a polynucleotide of (c); and

(e) a polynucleotide having at least 80% homology to the nucleotide sequence of SEQ ID No. 1.

The polynucleotide preferably encodes a polypeptide having *mpa* activity.

The invention also provides a polynucleotide probe or primer which comprises a fragment of at least 15 nucleotides of a polynucleotide selected from:

(b) a polynucleotide comprising a nucleotide sequence capable of hybridising to a fragment of the nucleotide sequence set out in SEQ ID No. 1, the fragment having the nucleotide sequence of nucleotides 210-1335 of SEQ ID No. 1;

(c) a polynucleotide comprising a nucleotide sequence capable of hybridising to the complement of a fragment of the nucleotide sequence set out in SEQ ID No. 1, the fragment having the nucleotide sequence of nucleotides 210-1335 of SEQ ID No. 1; and

(d') a polynucleotide comprising a polynucleotide sequence which is degenerate as a result of the genetic code to a polynucleotide sequence of (c).

The invention also provides a polypeptide in substantially isolated form which is encoded by a polynucleotide of the invention.

The invention also provides a polypeptide in substantially isolated form which comprises the sequence set out in SEQ ID No. 2, or a polypeptide substantially homologous thereto

which has *mpa* activity, or a fragment of the polypeptide of SEQ ID No. 2 which has *mpa* activity. The invention also provides a polypeptide comprising at least 8 amino acids which is an immunogenic fragment of said polypeptides and which comprises an epitope which is specific to the pathogenicity of mycobacterial cells.

The invention also provides a vector comprising a polynucleotide of the invention. Preferably the vector is an expression vector comprising a polynucleotide of the invention operably linked to regulatory sequences capable of directing expression of said polynucleotide in a host cell. The invention also provides an antibody capable of recognising a polypeptide of the invention. Preferably an antibody which is a monoclonal antibody or a fragment thereof.

The invention also provides a method for detecting the presence or absence of a polynucleotide of the invention in a biological sample which method comprises:

(a) bringing a biological sample containing DNA or RNA into contact with a probe of the invention under hybridising conditions; and

(b) detecting any duplex formed between the probe and nucleic acid in the sample.

The invention also provides a method of detecting the presence or absence of a polypeptide of the invention in a biological sample which method comprises:

(a) incubating the biological sample with an antibody of the invention under conditions which allow for the formation of an antibody-antigen complex; and

(b) determining whether antibody-antigen complex comprising said antibody is formed.

The invention also provides a method of detecting the presence or absence of antibodies in a biological sample which method comprises:

(a) incubating a biological sample with a polypeptide of the invention comprising an epitope under conditions which allow for the formation of an antibody-antigen complex; and

(b) determining whether an antibody-antigen complex comprising said polypeptide is formed.

The invention also provides a method of detecting the presence or absence of cell mediated

immune reactivity in an animal or human, to a peptide of the invention which method comprises:

(a) incubating a cell sample with a polypeptide of the invention comprising an epitope under conditions which allow for a cellular immune response; and

(b) detecting the presence of said cellular immune response in the incubate.

The invention also provides a test kit for detecting the presence or absence of a pathogenic mycobacterium in a sample which comprises a polynucleotide, a polypeptide or an antibody of the invention.

The invention also provides a pharmaceutical composition comprising (i) a polypeptide, a polynucleotide or an antibody of the invention and (ii) a suitable carrier or diluent.

The invention also provides a polypeptide, a polynucleotide or an antibody of the invention for use in the treatment, prevention or diagnosis of a disease caused by a mycobacterium.

The invention also provides a method of treating or preventing a mycobacterial disease in an animal or human caused by mycobacteria which express a polypeptide which method comprises administering to the animal or human an effective amount of said polypeptide.

The invention also provides a method of treating or preventing a mycobacterial disease in animals or humans caused by mycobacteria containing the nucleotide sequence of SEQ ID No. 1, which method comprises administering to the animal or human an effective amount of a polynucleotide or a vector of the invention.

Preferably the methods of treating mycobacterial disease of the invention are used against Johne's disease or Crohn's disease.

The invention also provides a method for increasing the *in vivo* susceptibility of mycobacteria to antimicrobial drugs.

The invention also provides a vaccine composition comprising (i) a polypeptide, a polynucleotide or a vector of the invention together with (ii) a pharmaceutically acceptable

carrier or diluent.

The invention also provides a plasmid containing a polynucleotide sequence of the invention under the control of a promoter.

The invention also provides a nucleic acid vaccine comprising (i) a plasmid of the invention; and (ii) a pharmaceutically acceptable carrier or diluent. Prefered nucleic acid vaccines which further comprise a transfection agent.

The invention also provides a vaccine comprising (i) a polypeptide of the invention optionally linked to a hapten molecule, and (ii) a pharmaceutically acceptable carrier or diluent.

The invention also provides a non-pathogenic microorganism or a cell from a human or animal species prone to infection by *mpa*-containing mycobacteria comprising a component on its surface which has been modified by a polypeptide of the invention.

The invention also provides a non-pathogenic microorganism or a cell from a human or animal species prone to infection by *mpa*-containing mycobacteria which has been transformed or transfected with a nucleic acid construct comprising a polynucleotide or plasmid, preferably the nucleic acid construct further comprises a polynucleotide which encodes the polypeptides of the GS region of *Mptb*. Most preferably the gene or genes present in the nucleic acid construct are expressed.

The invention also provides a vaccine comprising (i) a non-pathogenic microorganism or a cell from a human or animal of the invention and (ii) a pharmaceutically acceptable carrier or diluent.

The invention also provides a non-pathogenic microorganism or a cell from a human or animal species prone to infection by *mpa*-containing mycobacteria comprising on its surface an antigenic determinant capable of being produced by the action of a polypeptide of the invention and which is capable of eliciting antibodies which bind the surface of *Mptb*.

The invention also provides a normally pathogenic mycobacterium or pathogenic isolate thereof, whose pathogenicity is mediated in all or in part by the presence or expression of a polypeptide of the invention, which mycobacterium or isolate harbours an attenuating mutation in the polynucleotide sequence of the invention.

The invention also provides a vaccine comprising the mycobacterium or isolate of the invention and a pharmaceutically acceptable carrier or diluent. Preferably the attenuating mutation in the mycobacterium or isolate comprised in the vaccine is mediated by the insertion of one or more nucleotides.

The invention also provides a polynucleotide insertion element selected from:

(a) a polynucleotide comprising the nucleotide sequence set out in SEQ ID Nos. 3 or 4;

(b) a polynucleotide comprising a nucleotide sequence capable of hybridising to a fragment of the nucleotide sequence set out in SEQ ID No. 3, the fragment having the nucleotide sequence of nucleotides 1856-2543 of SEQ ID No. 3;

(c) a polynucleotide comprising a nucleotide sequence capable of hybridising to a fragment of the nucleotide sequence set out in SEQ ID No. 4, the fragment having the nucleotide sequence of nucleotides 1-688 of SEQ ID No. 4;

(d) a polynucleotide comprising a polynucleotide sequence which is degenerate as a result of the genetic code to the polynucleotide of SEQ ID No. 4 or a polynucleotide of (b);

(e) a polynucleotide having at least 75% homology to the nucleotide sequence of SEQ ID No. 3; and

(f) a polynucleotide having at least 75% homology to the nucleotide sequence of SEQ ID No. 4.

The invention provides as a preferred embodiment a vaccine comprising a mycobacterium or isolate of the invention wherein the attenuating mutation is mediated by insertion of a polynucleotide insertion element of the invention.

The invention also provides a polynucleotide probe or primer which comprises a fragment of at least 15 nucleotides of an insertion element of the invention, optionally carrying a revealing label.

The invention also provides a polypeptide in substantially isolated form which is encoded by an insertion element polynucleotide.

The invention also provides a polypeptide comprising at least 8 amino acids which is an immunogenic fragment of an insertion element polypeptide of the invention and which comprises an istA epitope.

The invention also provides a vector comprising an insertion element polynucleotide of the invention.

The invention also provides an expression vector comprising an insertion element polynucleotide of the invention, operably liked to regulatory sequences capable of directing expression of said polynucleotide in a host cell.

The invention also provides a method for preparing a mycobacterium or pathogenic isolate of the invention which method comprises transfecting animal or human isolate of an *mpa* containing pathogenic bacterium with a polynucleotide construct comprising an insertion element polynucleotide of the invention. In preferred embodiments transfection is effected by electroporation. In most preferred embodiments the polynucleotide comprised in the construct has the nucleotide sequence set out in SEQ ID No. 3 or 4.

Detailed description of the invention.

A. Polynucleotides.

Polynucleotides of the invention may comprise DNA or RNA. They may be single or double stranded. They may also be polynucleotides which include within them synthetic or modified nucleotides. Polynucleotides of the invention include polynucleotides in substantially isolated and isolated form. A number of different types of modification to oligonucleotides are known in the art. These include methylphosphonate and phosphorothioate backbones, addition of acridine or polylysine chains at the 3' and/or 5'

ends of the molecule. For the purposes of the present invention, it is to be understood that the polynucleotides described herein may be modified by any method available in the art. Such modifications may be carried out in order to enhance the *in vivo* activity or life span of polynucleotides of the invention.

Polynucleotides of the invention include:

i. a polynucleotide comprising the nucleotide sequence set out in SEQ ID No. 1 or the complement thereof;

ii. a polynucleotide comprising a polynucleotide sequence capable of hybridising to a polynucleotide having the sequence set out in SEQ ID No. 1 or a fragment thereof, preferred fragments having the sequence set out in nucleotides 110-1335 preferably 210-1335 of SEQ ID No. 1;

iii. a polynucleotide comprising a nucleotide sequence capable of hybridising to the complement of a polynucleotide having the sequence set out in SEQ ID No. 1 or a fragment thereof, preferred fragments having the nucleotide sequence set out in nucleotides 110-1335 preferably 210-1335 of SEQ ID No. 1;

iv. a polynucleotide comprising a polynucleotide sequence which is degenerate as a result of the genetic code to the nucleotide of SEQ ID No. 1 or a polynucleotide of (iii);

v. a polynucleotide having at least 80% homology to the nucleotide sequence of SEQ ID No. 1 or its complement;

vi. a polynucleotide comprising a nucleotide sequence set out in SEQ ID No. 3 or the complement thereof (SEQ ID No. 4);

vii. a polynucleotide comprising a nucleotide sequence capable of hybridising to a fragment of a polynucleotide having the sequence set out in SEQ ID No. 3, the

fragment having the nucleotide sequence set out in nucleotides 1856-2543 of SEQ ID No. 3;

viii. a polynucleotide comprising a nucleotide sequence capable of hybridising to a fragment of a polynucleotide having the sequence set out in SEQ ID No. 4, the fragment having the nucleotide sequence of nucleotides 1-688 of SEQ ID No. 4;

ix. a polynucleotide comprising a polynucleotide sequence which is degenerate as a result of the genetic code to the polynucleotide of SEQ ID No. 4 or the polynucleotide of (vii);

x. a polynucleotide having at least 75% homology to the nucleotide sequence of SEQ ID No. 3; and

xi. a polynucleotide having at least 75% homology to the nucleotide sequence of SEQ ID No. 4.

Polynucleotides of the invention which are described as capable of hybridising to all or part of the DNA of SEQ ID No. 1 or the complements thereof or fragments of SEQ ID Nos. 3 and 4 will be generally at least 70%, preferably at least 75%, 80% or 90% or more preferably at least 95% homologous to the DNA to which they are described as hybridising over a region of at least 20, preferably at least 25 or 30 for instance at least 40, 60 or 100 or more contiguous nucleotides.

It is to be understood that skilled persons may, using routine techniques, make nucleotide substitutions that do not affect the polypeptide sequence encoded by the polynucleotides of the invention to reflect the codon usage of any particular host organism in which the polypeptides of the invention are to be expressed.

Any combination of the above mentioned degrees of homology and minimum sizes may be used to define polynucleotides of the invention, with the more stringent combinations (i.e. higher homology over longer lengths) being preferred. Thus for example a polynucleotide which is at least 80% homologous over 25, preferably 30 nucleotides forms one aspect of the invention, as does a polynucleotide which is at least 90% homologous over 40 nucleotides.

Polynucleotides of the invention may be used to produce a primer, e.g. a PCR primer, a primer for an alternative amplification reaction, a probe e.g. labelled with a revealing label by conventional means using radioactive or non-radioactive labels, or the polynucleotides may be cloned into vectors. Such primers, probes and other fragments will be at least 15, preferably at least 20, for example at least 25, 30 or 40 nucleotides in length, and are also enco*mpas*sed by the term polynucleotides of the invention as used herein.

Polynucleotides such as a DNA polynucleotide and primers according to the invention may be produced recombinantly, synthetically, or by any means available to those of skill in the art. They may also be cloned by standard techniques.

In general, primers will be produced by synthetic means, involving a step wise manufacture of the desired nucleic acid sequence one nucleotide at a time. Techniques for accomplishing this using automated techniques are readily available in the art.

Longer polynucleotides will generally be produced using recombinant means, for example using a PCR (polymerase chain reaction) cloning techniques. This will involve making a pair of primers (e.g. of about 15-30 nucleotides) to a region of the *mpa* gene which it is desired to clone, bringing the primers into contact with mRNA or cDNA, performing a

polymerase chain reaction under conditions which bring about amplification of the desired region, isolating the amplified fragment (e.g. by purifying the reaction mixture on an agarose gel) and recovering the amplified DNA. The primers may be designed to contain suitable restriction enzyme recognition sites so that the amplified DNA can be cloned into a suitable cloning vector. Although in general the techniques mentioned herein are well known in the art, reference may be made in particular to Sambrook *et al.*, Molecular Cloning, A Laboratory Manual (1989) and Ausubel *et al.*, Current Protocols in Molecular Biology (1995), John Wiley & Sons, Inc.

Polynucleotides which are not 100% homologous to the sequences of the present invention but fall within the scope of the invention can be obtained in a number of ways. Other isolates or strains of pathogenic mycobacteria will be expected to contain allelic variants of the *mpa* sequence described herein may be obtained for example by probing genomic DNA libraries made from such pathogenic mycobacteria. In addition, other mycobacterial homologues of *mpa* in pathogenic *M.avium* may be obtained and such homologues and fragments thereof in general will be capable of selectively hybridising to SEQ ID No. 1. Such sequences may be obtained by probing cDNA libraries made from or genomic DNA libraries from other mycobacterial species and their isolates, and probing such libraries with probes comprising all or part of SEQ ID. No. 1 as defined above under conditions of medium to high stringency (for example 0.03M sodium chloride and 0.03M sodium citrate at from about 50°C to about 60°C). Homologues of the *mpa* gene in other pathogenic *M.avium* may also be obtained by polymerase chain reaction or other method of amplifying these *mpa* genes using primers derived from SEQ ID No.1 or the complement thereof.

Allelic variants and species homologues may also be obtained using degenerate PCR which

will use primers designed to target sequences within the variants and homologues encoding conserved amino acid sequences. The primers will contain one or more degenerate positions and will be used at stringency conditions lower than those used for cloning sequences with single sequence primers against known sequences.

Alternatively, such polynucleotides may be obtained by site directed mutagenesis of the *mpa* sequence or allelic variants thereof. This may be useful where for example any silent codon changes are required to sequences in order to optimise codon preferences for a particular host cell in which the polynucleotide sequences are being expressed. Other sequence changes may be desired in order to introduce restriction enzyme recognition sites, or to alter the property or function of the polypeptides encoded by the polynucleotides.

Such altered property or function will include the addition of amino acid sequences of consensus signal peptides known in the art to effect transport and secretion of the modified polypeptide of the invention. Another altered property will include mutagenesis of a catalytic residue or generation of fusion proteins with another polypeptide. Such fusion proteins may be with an enzyme, with an antibody or with a cytokine or other ligand for a receptor, to target a polypeptide of the invention to a specific cell type *in vitro* or *in vivo* or to enhance immune recognition.

The invention further provides double stranded polynucleotides comprising a polynucleotide of the invention and its complement.

Polynucleotides or primers of the invention may carry a revealing label. Suitable labels include radioisotopes such as ^{32}P or ^{35}S, enzyme labels, or other protein labels such as

biotin. Such labels may be added to polynucleotides or primers of the invention and may be detected using by techniques known *per se*.

Polynucleotides or primers of the invention or fragments thereof labelled or unlabelled may be used by a person skilled in the art in nucleic acid-based tests for detecting or sequencing *mpa* and its homologues in samples of body fluids, tissues or excreta from animals and humans and to food and environmental samples e.g. water. Human and animal body fluids include sputum, blood, serum, plasma, saliva, milk, urine, csf, semen, faeces and infected discharges. Tissues include intestine, mouth ulcers, skin, lymph nodes, spleen, lung and liver obtained surgically or by a biopsy technique. Animals particularly include commercial livestock such as cattle, sheep, goats, deer, rabbits but wild animals and animals in zoos may also be tested.

Such tests for detecting generally comprise bringing a biological sample containing DNA or RNA into contact with a probe comprising a polynucleotide or primer of the invention under hybridising conditions and detecting any duplex formed between the probe and nucleic acid in the sample. Such detection may be achieved using techniques such as PCR or by immobilising the probe on a solid support, removing nucleic acid in the sample which is not hybridised to the probe, and then detecting nucleic acid which has hybridised to the probe. Alternatively, the sample nucleic acid may be immobilised on a solid support, and the amount of probe bound to such a support can be detected. Suitable assay methods of this any other formats can be found in for example WO89/03891 and WO90/13667.

Methods for sequencing *mpa* or IS*1612* and their homologues include bringing a biological sample containing target DNA or RNA into contact with a probe comprising a

polynucleotide or primer of the invention under hybridising conditions and determining the sequence by, for example the Sanger dideoxy chain termination method (see Sambrook *et al.*). Such a method generally comprises elongating, in the presence of suitable reagents, the primer by synthesis of a strand are to the target DNA or RNA and selectively terminating the elongation reaction at one or more of an A, C, G or T/U residue; allowing strand elongation and termination reaction to occur; separating out according to size the elongated products to determine the sequence of the nucleotides at which selective termination has occurred. Suitable reagents include a DNA polymerase enzyme, the deoxynucleotides dATP, dCTP, dGTP and dTTP, a buffer and ATP. Dideoxynucleotides are used for selective termination.

Polynucleotides of the invention or fragments thereof labelled or unlabelled may also be used to identify and characterise different strains of *Mptb*, and other *mpa*-containing pathogenic mycobacteria and properties such as drug resistance or susceptibility. The probes of the invention may conveniently be packaged in the form of a test kit in a suitable container for storage and transport. In such kits the probe may be bound to a solid support where the assay format for which the kit is designed requires such binding. The kit may also contain suitable reagents for treating the sample to be probed, hybridising the probe to nucleic acid in the sample, control reagents, instructions, and the like. The kit also preferably involves methods known in the art such as PCR or LCR.

The use of polynucleotides of the invention in the diagnosis of inflammatory diseases such as Crohn's disease or sarcoidosis in humans or Johne's disease in animals form preferred aspects of the invention. The polynucleotides may also be used in the prognosis of these diseases. For example, the response of a human or animal subject in response to antibiotic,

vaccination or other therapies may be monitored by utilising the diagnostic methods of the invention over the course of a period of treatment and following such treatment.

The use of *mpa* polynucleotides (particularly in the form of probes and primers) of the invention in the above-described methods form a further aspect of the invention, particularly for the detection, diagnosis or prognosis of *Mptb* infections.

The present invention also provides polynucleotides encoding the polypeptides of the invention described below. Because such polynucleotides will be useful as sequences for recombinant production of polypeptides of the invention, it is not necessary for them to be selectively hybridisable to the sequence of SEQ ID No. 1 although this will generally be desirable. Otherwise, such polynucleotides may be labelled, used, and made as described above if desired. Preferably the polynucleotides of the invention encode polypeptides which possess *mpa* activity. *Mpa* polypeptides of the invention are described below.

The present invention also provides the polynucleotide sequence of an IS*21*-like insertion element hereafter termed IS*1612* and related polynucleotide sequences as described in items vi-xi and relevant passages above (hereafter referred to as the IS*1612* polynucleotides of the invention). The invention in this aspect also provides a polynucleotide construct comprising the IS*1612* polynucleotides of the invention which construct is in a suitable form for transfection into *mpa*-containing bacterial cells, preferably *mpa*-containing mycobacterial cells originating from an animal or human. The invention further provides a method for attenuating a normally pathogenic mycobacterium whose pathogenicity is mediated in all or in part by the presence or expression of the *mpa* polynucleotides (i-v and relevant passages above) of the invention, which method comprises transfecting into a mycobacterial cell or

isolate of said pathogenic mycobacterium a polynucleotide construct comprising the IS*1612* polynucleotides of the invention.

Transfection can be carried out by electroporation or any other method known in the art such that the function of the *mpa* gene is knocked out, resulting in a mutated attenuated form of these normally pathogenic mycobacteria.

The IS*1612* element encodes from SEQ ID No. 4 two polypeptides the first of which in its entirity is new and forms part of the present invention. The entire amino acid sequence of the first polypeptide, termed *ist*A, is encoded by nucleotides 98-1678 of SEQ ID No. 4 and is given as SEQ ID No. 5. The IS*1612* polypeptides from the invention are further described below.

B. Polypeptides.

Polypeptides of the invention include polypeptides in substantially isolated form which comprise the sequence set out in SEQ ID No. 2 or SEQ ID No. 5. Polypeptides further include variants of such sequences, including naturally occurring variants and synthetic variants which are substantially homologous to said polypeptides. In this context, substantial homology is regarded as a sequence which has at least 70%, e.g. 80%, 90%, 95% or 98% amino acid homology (identity) over 30 amino acids with the sequence of SEQ ID No. 2.

Polypeptides of the invention also include fragments of the above mentioned full length polypeptides and variants thereof, including in particular any fragments of the sequence set

out in SEQ ID No. 2 and the fragment having the sequence of amino acids 1-211 as set out in SEQ ID No. 5. Preferred fragments include those which include an epitope and especially in the use of related polypeptides to SEQ ID No. 2 an epitope which is specific to the pathogenicity of the mycobacterial cell from which the polypeptide is derived. Suitable fragments will be at least 8, e.g. 10, 12, 15 or 20 amino acids in size. Polypeptide fragments of *mpa* and allelic and species variants thereof may contain one or more (e.g. 2, 3, 5, or 10) substitutions, deletions or insertions, including conserved substitutions, or modifications of amino acids such as phosphorylation of serine or threonine.

Conserved substitutions may be made according to the following table which indicates conservative substitutions, where amino acids on the same block in the second column and preferably in the same line in the third column may be substituted for each other:

ALIPHATIC	Non-polar	G A P
		I L V
	Polar – uncharged	C S T M
		N Q
	Polar – charged	D E
		K R
AROMATIC		H F W Y
OTHER		N Q D E

Epitopes may be determined, for example, by techniques such as peptide scanning techniques as described by Geysen *et al*, Mol. Immunol., 23, 709-715 (1986) as well as other techniques known in the art.

The term "an epitope which is specific to the pathogenicity of the mycobacterial cell" means that the epitope is encoded by a portion of the *mpa* gene, or by the corresponding ORF sequences in other mycobacteria which can be used to distinguish mycobacteria which are pathogenic from related non-pathogenic mycobacteria including non-pathogenic species of *M.avium*. This may be determined using routine methodology. A candidate epitope may be prepared and used to immunise an animal such as a rat or rabbit in order to generate antibodies. The antibodies may then be used to detect the presence of the epitope in pathogenic mycobacteria and to confirm that non-pathogenic mycobacteria do not contain any proteins which react with the epitope. Epitopes may be linear or conformational.

Polypeptides of the invention may be in a substantially isolated form. It will be understood that the polypeptide may be mixed with carriers or diluents which will not interfere with the intended purpose of the polypeptide and still be regarded as substantially isolated. A polypeptide of the invention may also be in a substantially purified form, in which case it will generally comprise the polypeptide in a preparation in which more than 90%, e.g. 95%, 98% or 99% of the polypeptide in the preparation is a polypeptide of the invention. Polypeptides of the invention may be modified for example by the addition of histidine residues to assist their purification or by the addition of a signal sequence to promote their secretion from a cell, or a transmembrane sequence to facilitate their immobilisation in a membrane or lipid bilayer.

Thus, polypeptides of the invention include in particular fusion proteins which comprise a polypeptide encoded all or part of the *mpa* gene fused at its N- or C-terminus to a second sequence to provide the desired property or function. Sequences which promote secretion from a cell include, for example the yeast α-factor signal sequence.

A polypeptide of the invention may be labelled with a revealing label. The revealing label may be any suitable label which allows the polypeptide to be detected. Suitable labels include radioisotopes e.g. ^{125}I, enzymes, antibodies, polynucleotides and linkers such as biotin. Labelled polypeptides of the invention may be used in diagnostic procedures such as immunoassays in order to determine the amount of a polypeptide of the invention in a sample. Polypeptides or labelled polypeptides of the invention may also be used in serological or cell mediated immune assays for the detection of immune reactivity to said polypeptides in animals and humans using standard protocols.

A polypeptide or labelled polypeptide of the invention or fragment thereof may also be fixed to a solid phase, for example the surface of an immunoassay well or dipstick. Such labelled and/or immobilised polypeptides may be packaged into kits in a suitable container for transport and storage along with suitable reagents, controls, instructions and the like.

Such polypeptides and kits may be used in methods of detection of antibodies or cell mediated immunoreactivity, to the mycobacterial proteins and peptides encoded by the *mpa* gene and its allelic variants and fragments, using immunoassay. Such host antibodies or cell mediated immune reactivity will occur in humans or animals with an immune system which detects and reacts against polypeptides of the invention. The antibodies may be present in a biological sample from humans or animals. The biological sample may be a sample as defined above particularly blood, milk or saliva.

Immunoassay methods are well known in the art and will generally comprise:

(a) providing a polypeptide of the invention comprising an epitope bindable by an antibody against said mycobacterial polypeptide;

(b) incubating a biological sample with said polypeptide under conditions which allow for the formation of an antibody-antigen complex; and

(c) determining whether antibody-antigen complex comprising said polypeptide is formed.

Immunoassay methods for cell mediated immune reactivity in animals and humans are also well known in the art (e.g. as described by Weir et al 1994, J.Immunol Methods 176; 93-101) and will generally comprise:

(a) providing a polypeptide of the invention comprising an epitope bindable by a lymphocyte or macrophage or other cell receptor;

(b) incubating a cell sample with said polypeptide under conditions which allow for a cellular immune response such as release of cytokines or other mediator to occur; and

(c) detecting the presence of said cytokine or mediator in the incubate.

Polypeptides of the invention may be made by synthetic means (e.g. as described by Geysen *et al.*, 1996) or recombinantly, as described below.

Particularly preferred polypeptides of the invention include Sequence ID No.2 and amino acids 1-211 of Sequence ID No.5. Fragments as defined above from these regions are particularly preferred. The polypeptides and fragments thereof may contain amino acid alterations as defined above.

Polypeptides of the invention or fragments thereof labelled or unlabelled may also be used to identify and characterise different strains of *Mptb,* or other *mpa*-containing pathogenic

mycobacteria and properties such as drug resistance or susceptibility.

The use of polypeptides of the invention in the diagnosis of inflammatory diseases such as Crohn's disease or sarcoidosis in humans or Johne's disease in animals form a preferred aspect of the invention. The polypeptides may also be used in the prognosis of these diseases. For example, the response of a human or animal subject to antibiotic or other therapies may be monitored by utilizing the diagnostic methods of the invention over the course of a period of treatment and following such treatment. The use of *mpa* polypeptides of the invention in the above-described methods form a further aspect of the invention, particularly for the detection, diagnosis or prognosis of mycobacterial preferably *Mptb* infections.

Polypeptides of the invention may also be used in assay methods for identifying candidate chemical compounds which will be useful in inhibiting, binding to or disrupting the function of said polypeptides required for pathogenicity. In general, such assays involve bringing the polypeptide into contact with a candidate inhibitor compound and observing the ability of the compound to disrupt, bind to or interfere with the polypeptide. There are a number of ways in which the assay may be formatted. For example, those polypeptides which have an enzymatic function may be assayed using labelled substrates for the enzyme, and the amount of, or rate of, conversion of the substrate into a product measured, e.g. by chromatography such as HPLC or by a colourimetric assay. Suitable labels include ^{35}S, ^{125}I, biotin or enzymes such as horse radish peroxidase.

Candidate chemical compounds which may be used may be natural or synthetic chemical compounds used in drug screening programmes. Extracts of plants which contain several

characterised or uncharacterised components may also be used.

Alternatively, a polypeptide of the invention may be screened against a panel of peptides, nucleic acids or other chemical functionalities which are generated by combinatorial chemistry. This will allow the definition of chemical entities which bind to polypeptides of the invention. Typically, the polypeptide of the invention will be brought into contact with a panel of compounds from a combinatorial library, with either the panel or the polypeptide being immobilized on a solid phase, under conditions suitable for the polypeptide to bind to the panel. The solid phase will then be washed under conditions in which only specific interactions between the polypeptide and individual members of the panel are retained, and those specific members may be utilised in further assays or used to design further panels of candidate compounds. Binding of polypeptides of the invention to specific ligands may be identified using an antibody to said polypeptide or a polypeptide bearing a revealing label.

For example, a number of assay methods to define peptide interaction with peptides are known. For example, WO86/00991 describes a method for determining mimotopes which comprises making panels of catamer preparations, for example octamers of amino acids, at which one or more of the positions is defined and the remaining positions are randomly made up of other amino acids, determining which catamer binds to a protein of interest and re-screening the protein of interest against a further panel based on the most reactive catamer in which one or more additional designated positions are systematically varied. This may be repeated throughout a number of cycles and used to build up a sequence of a binding candidate compound of interest.

WO89/03430 describes screening methods which permit the preparation of specific

mimotopes which mimic the immunological activity of a desired analyte. These mimotopes are identified by reacting a panel of individual peptides wherein said peptides are of systematically varying hydrophobicity, amphipathic characteristics and charge patterns, using an antibody against an antigen of interest. Thus in the present case antibodies against a polypeptide of the invention may be employed and mimotope peptides from such panels may be identified. We particularly recognise that antibodies, including monoclonal antibodies, to GPL's from pathogenic *Mptb* and other pathogenic *M.avium* species containing an intact functional *mpa* gene may be identified by their ability to bind to these organisms, but not to mutated forms in which the function of the *mpa* gene has been knocked-out. These antibodies may be used to screen phage-displayed peptide libraries for immunogenic mimics of GPL epitopes, and that such peptide immunogenic mimics may be used in the development of anti-GPL vaccines [Phalipon et al. Eur.J.Immunol. 27:2620-2625, 1997].

C. Vectors.

Polynucleotides of the invention can be incorporated into a recombinant replicable vector. The vector may be used to replicate the nucleic acid in a compatible host cell. Thus in a further embodiment, the invention provides a method of making polynucleotides of the invention by introducing a polynucleotide of the invention into a replicable vector, introducing the vector into a compatible host cell, and growing the host cell under conditions which bring about replication of the vector. The vector may be recovered from the host cell. Suitable host cells are described below in connection with expression vectors.

D. Expression Vectors.

Preferably, a polynucleotide of the invention in a vector is operably linked to a regulatory sequence which is capable of providing for the expression of the coding sequence by the host cell, i.e. the vector is an expression vector. The term "operably linked" refers to a juxtaposition wherein the components described are in a relationship permitting them to function in their intended manner. A regulatory sequence "operably linked" to a coding sequence is ligated in such a way that expression of the coding sequence is achieved under condition *compa*tible with the control sequences.

The term "regulatory sequences includes promoters and enhancers and other expression regulation signals. These may be selected to be co*mpa*tible with the host cell for which the expression vector is designed. For example, yeast regulatory sequences include *S. cerevisiae* GAL4 and ADH promoters, *S. pombe* nmt1 and adh promoters. Mammalian promoters, such as -actin promoters, may be used. Mammalian promoters also include the metallothionein promoter which can upregulate expression in response to heavy metals such as cadmium and is thus an inducible promoter. Tissue-specific promoters, for example neuronal cell specific may be used. Viral promoters may also be used, for example the Moloney murine leukaemia virus long terminal repeat (MMLV LTR), the promoter rous sarcoma virus (RSV) LTR promoter, the SV40 promoter, the human cytomegalovirus (CMV) IE promoter, herpes simplex virus promoters or adenovirus promoters. All these promoters are readily available in the art.

Such vectors may be transformed into a suitable host cell as described above to provide for expression of a polypeptide of the invention. Thus, in a further aspect the invention

provides a process for preparing polypeptides according to the invention which comprises cultivating a host cell transformed or transfected with an expression vector as described above under conditions to provide for expression by the vector of a coding sequence encoding the polypeptides, and recovering the expressed polypeptides.

The vectors may be for example, plasmid, virus or phage vectors provided with an origin of replication, optionally a promoter for the expression of the said polynucleotide and optionally a regulator of the promoter. The vectors may contain one or more selectable marker genes, for example an ampicillin resistance gene in the case of a bacterial plasmid or a neomycin resistance gene for a mammalian vector. Vectors may be used *in vitro*, for example for the production of RNA or used to transfect or transform a host cell. The vector may also be adapted to be used *in vivo*, for example in a method of gene therapy or a DNA vaccine.

The use of mammalian host cells is expected to provide for such post-translational modifications (e.g. myristolation, glycosylation, truncation, lapidation and tyrosine, serine or threonine phosphorylation) as may be needed to confer optimal biological activity on recombinant expression products of the invention. Such cell culture systems in which polypeptide of the invention are expressed may be used in assay systems to identify candidate substances which interfere with or enhance the functions of the polypeptides of the invention in the cell.

A further embodiment of the invention provides host cells transformed or transfected with the vectors for the replication and expression of polynucleotides of the invention. The cells will be chosen to be co*mpa*tible with the said vector and may for example be bacterial,

yeast, insect or mammalian.

Polynucleotides according to the invention may also be inserted into the vectors described above in an antisense orientation in order to provide for the production of antisense RNA. Antisense RNA or other antisense polynucleotides may also be produced by synthetic means. Such antisense polynucleotides may be used in a method of controlling the levels of *mpa* or its variants or species homologues.

E. Antibodies.

The invention also provides monoclonal or polyclonal antibodies to polypeptides of the invention or fragments thereof. The invention further provides a process for the production of monoclonal or polyclonal antibodies to polypeptides of the invention. Monoclonal antibodies may be prepared by conventional hybridoma technology using the polypeptides of the invention or peptide fragments thereof, as immunogens. Polyclonal antibodies may also be prepared by conventional means which comprise inoculating a host animal, for example a rat or a rabbit, with a polypeptide of the invention or peptide fragment thereof and recovering immune serum. In order that such antibodies may be made, the invention also provides polypeptides of the invention or fragments thereof haptenised to another polypeptide for use as immunogens in animals or humans.

For the purposes of this invention, the term "antibody", unless specified to the contrary, includes fragments of whole antibodies which retain their binding activity for a tumour target antigen. Such fragments include Fv, F(ab') and $F(ab')_2$ fragments, as well as single chain antibodies. Furthermore, the antibodies and fragments thereof may be humanised antibodies, e.g. as described in EP-A-239400.

Antibodies may be used in method of detecting polypeptides of the invention present in biological samples by a method which comprises:

(a) providing an antibody of the invention;

(b) incubating a biological sample with said antibody under conditions which allow for the formation of an antibody-antigen complex; and

(c) determining whether antibody-antigen complex comprising said antibody is formed.

Antibodies of the invention may be bound to a solid support and/or packaged into kits in a suitable container for storage and transport along with suitable reagents, controls, instructions and the like. Antibodies of the invention may be used in the detection, diagnosis and prognosis of diseases as discussed above in relation to polypeptides of the invention.

F. Compositions.

The present invention also provides compositions comprising a polynucleotide or polypeptide of the invention together with a carrier or diluent. Compositions of the invention also include compositions comprising a nucleic acid, particularly and expression vector, of the invention. Compositions further include those carrying a recombinant virus of the invention. Such compositions include pharmaceutical compositions in which case the carrier or diluent will be pharmaceutically acceptable.

Pharmaceutically acceptable carriers or diluents include those used in formulations suitable for inhalation as well as oral, parenteral (e.g. intramuscular or intravenous or transcutaneous) administration. The formulations may conveniently be presented in unit

dosage form and may be prepared by any of the methods well known in the art of pharmacy. Such methods include the step of bringing into association the active ingredient with the carrier which constitutes one or more accessory ingredients. In general the formulations are prepared by uniformly and intimately bringing into association the active ingredient with liquid carriers or finely divided solid carriers or both, and then, if necessary, shaping the product.

For example, formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostatis and solutes which render the formulation isotonic with the blood of the intended recipient, and aqueous and non-aqueous sterile suspensions which may include suspending agents and thickening agents, and liposomes or other microparticulate systems such as gold particles for use in the DNA gun, all of which are designed to target the polynucleotide or the polypeptide of the invention to blood components or one or more organs, or to target cells such as skin cells, muscle cells, or mucosal M cells of the intestine after oral administration or rectal administration or of the nasal and respiratory mucosa.

G. Vaccines.

In another aspect, the invention provides novel vaccines for the prevention and treatment of infections caused by *Mptb* and other *mpa*-containing pathogenic mycobacteria in animals and humans. The term "vaccine" as used herein means an agent used to stimulate the immune system of a vertebrate, particularly a warm blooded vertebrate including humans, so as to provide protection against future harm by an organism to which the vaccine is directed or to assist in the eradication of an organism in the treatment of established

infection. The immune system will be stimulated by the production of cellular immunity and antibodies, desirably neutralizing antibodies, directed to epitopes found on or in a pathogenic mycobacterium which expresses the *mpa* gene of the invention. The antibody so produced may be any of the immunological classes, such as the immunoglobulins A, D, E, G or M. Vaccines which stimulate the production of IgA are interest since this is the principle immunoglobulin produced by the secretory system of warm-blooded animals, and the production of such antibodies will help prevent infection or colonization of the intestinal tract. However an IgM and IgG response will also be desirable for vaccination against systemic infections such as Crohn's disease or Johne's disease.

Vaccines of the invention include naked nucleic acid vaccines such as DNA vaccines. Naked nucleic acid vaccines of the invention include polynucleotides of the invention or fragments thereof in suitable vectors which may be administered as naked nucleic acid using standard protocols. The preferred nucleic acid vaccines of the invention are DNA vaccines which comprise at least one DNA polynucleotide of the invention incorporated into a plasmid under the control of a strong promoter. Suitable strong promoters are viral promoters such as the Muloney murine leukaemia virus long-terminal repeat (MMLV LTR), the promoter rouse sarcoma virus (RSV) LTR promoter, the SV40 promoter, the human cytomegalovirus (CMV) IE promoter, herpes simplex virus promoters or adenovirus promoters. The preferred promoter is CMV. The nucleic acid vaccines of the invention made be administered intradermally, subcutaneously, intramuscularly or mucosally by inhalation, ingestion or intra-rectal installation. The nucleic acid vaccine of the invention may be incorporated into a suitable formulation, absorbed or coated onto particles which are appropriate for ingestion into macrophages or antigen presenting cells. Such particles may include gold particles and may be administered by using a gene gun. The nucleic acid

vaccine of the invention may also be carried within another organism such as disabled *salmonella*, *M.Bovis BCG, M. Smegmatis* or other mycobacteria, Corynebacteria, or other agents according to established protocol.

Immunising formulations may advantageously include an adjuvant.

When the polynucleotide of the invention is administered as a nucleic acid vaccine, the amount of nucleic acid administered is typically in the range of from 50-500 μg. Uptake of the nucleic acid vaccines of the invention by mammalian cells may be enhanced by several transfection techniques, for example, those including the use of transfection agents. Examples of these agents include catonication (for example calcium phosphate and DEAE-dextran) and lipofectants (for example lipofectam ™ and transfectam ™). Confirmation of successful expression of *mpa* can be achieved by analysing surface sugars for acetylation using high pressure liquid chromatography and other methods known in the art. Typically, the nucleic acid in the vaccine is mixed with the transfection agent to produce a composition. Preferably the nucleic acid vaccine with or without transfection agent is combined with a pharmaceutically acceptable carrier or diluent to produce a pharmaceutical composition. Suitable carriers and diluents include isotonic saline solution, for example, phosphate buffered saline. The route of administration and dosages described are intended only a guide since the skilled practitioner will be able to readily determine the optimum route of administration and dosage for any particular patient and condition.

In addition to the above it is also possible to prepare "live" vaccines containing non-pathogenic microorganisms which express or are capable of causing to be expressed one or more polypeptides of the invention. Non-pathogenic microorganisms may be naturally non-

pathogenic microorganisms or attenuated microorganisms which are normally pathogenic. Suitable attenuated microorganisms are known in the art and include attenuated bacteria such as attenuated salmonella and attenuated viruses such as vaccinia virus. Transformation of the *mpa* gene or other polynucleotides of the invention into the attenuated microorganism is preferably completed *in vitro*. It is also within the invention to transfect cells of a human or animal to be vaccinated with a polynucleotide of the invention *in vitro* although it is preferred that transfection of a host cell should occur *in vivo* by, for example, DNA vaccination as described above.

It shall also be appreciated that the invention provides a non-pathogenic microorganism or human or animal cell which has on its surface a cell surface component produced by modification of a cell surface substrate with a polypeptide of the invention. Preferably, the polypeptide of the invention is expressed within the non-pathogenic microorganism or human or animal cell upon which the modified cell surface component resides. Thus the invention provides a non-pathogenic microorganism or human or animal cell which has been transformed or transfected with a polynucleotide of the invention which is expressed and results in the presence of the modified cell surface component. It should also be appreciated that it is within this invention to treat the surface of cells exhibiting the cell surface substrate with preformed polypeptide of the invention to produce cells having the modified cell surface component. It is also within this invention to provide a non-pathogenic microorganism or human or animal cell wherein an identical antigenic determinant to the modified cell surface component is produced on a cell surface using polypeptides other than *mpa* which fulfil an identical role to *mpa*.

In a further aspect vaccines of the invention which function by expressing *mpa* including,

for example, nucleic acid vaccines or live vaccines mentioned above, may be augmented by expressing all of the polypeptides encoded for in the ORFs of the GS region identified in *Mptb* along with *mpa* of the invention. The polypeptides encoded for by the ORFs of GS in *Mptb* and the polynucleotide sequence of GS in *Mptb* were disclosed and characterised in PCT/GB96/03221. These vaccines may be obtained by transforming the attenuated microorganism mentioned above with a polynucleotide of the invention which comprises the *mpa* gene or previously defined related sequence thereof together with the genes encoding the polypeptides associated with the GS region in *Mptb*. For the naked nucleic acid vaccines transfection of a host cell *in vivo* may be achieved by nucleic acid vaccination using the *mpa* gene or the previously defined related sequence thereof together with the genes encoding the polypeptides associated with the GS element of *Mptb* under the control of one or more suitable promoters.

Polypeptides of the invention or fragments thereof in substantially isolated form may also be used as vaccines by injection, inhalation, oral administration or by transcutaneous application according to standard protocols. Adjuvants (such as Iscoms or polylactide-coglycolide encapsulation), cytokines such as IL-12 and other immunomodulators may be used for the selective enhancement of the cell mediated or humeral immunological responses. Vaccination with polynucleotides and/or polypeptides of the invention may be undertaken to increase the susceptibility of pathogenic mycobacteria to antimicrobial agents *in vivo*.

In instances wherein the polypeptide is correctly configured so as to provide the correct epitope, but is too small to be immunogenic, the polypeptide may be linked to a suitable carrier.

A number of techniques for obtaining such linkage are known in the art, including the formation of disulphide linkages using N-succinimidyl-3-(2-pyridylthio) propionate (SPDP) and succinimidyl 4-(N-maleimido-methyl)cyclohexane-1-carboxylate (SMCC) obtained from Pierce Company, Rockford, Illinois, (if the peptide lacks a sulfhydryl group, this can be provided by addition of a cysteine residue). These reagents create a disulphide linkage between themselves and peptide cysteine residues on one protein and an amide linkage through the epsilon-amino on a lysine, or other free amino group in the other. A variety of such disulphide/amide-forming agents are known. See, for example, Immun Rev (1982) 62:185. Other bifunctional coupling agents form a thioether rather than a disulphide linkage. Many of these thio-ether-forming agents are commercially available and include reactive esters of 6-maleimidocaproic acid, 2-bromoacetic acid, 2-iodoacetic acid, 4-(N-maleimido-methyl) cyclohexane-1-carboxylic acid, and the like. The carboxyl group can be activated by combining them with succinimide or 1-hydroxyl-2-nitro-4-sulfonic acid, sodium salt. Additional methods of coupling antigens employs the rotavirus/"binding peptide" system described in EP 259,149, the disclosure of which is incorporated herein by reference. The foregoing list is not meant to be exhaustive, and modifications of the named compounds can clearly be used.

Any carrier may be used which does not itself induce the production of antibodies harmful to the host. Suitable carriers are typically large, slowly metabolized macromolecules such as proteins; polysaccharides, such as latex functionalized Sepharose 7, agarose, cellulose, cellulose beads and the like; polymeric amino acids, such as polyglutamic acid and polylysine, and the like; amino acid copolymers; and inactive virus particles. Especially useful protein carriers are serum albumens, keyhole limpet hemocyanin, immunoglobulin molecules, thyroglobulin, ovalbumin, tetanus toxoid, and other proteins well known to

those skilled in the art. Particularly useful also are controlled release microspheres composed of polylactide coglycolide typically 75:25 and typically 50μM in diameter. Polynucleotide and polypeptide vaccines of the invention adsorbed onto such microspheres may be zinc stabilised. Lamella polymers of chitin or chitosan in microparticulate form typically 100 nM may also be employed.

The immunogenicity of the epitopes may also be enhanced by preparing them in mammalian or yeast systems fused with or assembled with particle-forming proteins such as, for example, that associated with hepatitis B surface antigen. See, e.g., US-A-4,722,840. Constructs wherein the epitope is linked directly to the particle-forming protein coding sequences produce hybrids which are immunogenic with respect to the epitope. In addition, all of the vectors prepared include epitopes specific to HBV, having various degrees of immunogenicity, such as, for example, the pre-S peptide.

In addition, portions of the particle-forming protein coding sequence may be replaced with codons encoding an epitope of the invention. In this replacement, regions which are not required to mediate the aggregation of the units to form immunogenic particles in yeast or mammals can be deleted, thus eliminating additional HBV antigenic sites from competition with the epitope of the invention.

Vaccines may be prepared from one or more immunogenic polypeptides of the invention.

The polypeptides for use in vaccines may be expressed in various host cells (e.g., bacteria, yeast, insect, or mammalian cells), or alternatively may be isolated from viral preparations or made synthetically.

The preparation of vaccines which contain an immunogenic polypeptide(s) as active ingredients, is known to one skilled in the art. Typically, such vaccines are prepared as injectables, or as suitably encapsulated oral preparations and either liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid prior to ingestion or injection may also be prepared. The preparation may also be emulsified, or the protein encapsulated in liposomes. The active immunogenic ingredients are often mixed with excipients which are pharmaceutically acceptable and co*mpa*tible with the active ingredient. Suitable excipients are, for example, water, saline, dextrose, glycerol, ethanol, or the like and combinations thereof. In addition, if desired, the vaccine may contain minor amounts of auxiliary substances such as wetting or emulsifying agents, pH buffering agents, and/or adjuvants which enhance the effectiveness of the vaccine. Examples of adjuvants which may be effective include but are not limited to: aluminum hydroxide, N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP), N-acetyl-nor-muramyl-L-alanyl-D-isoglutamine (CGP 11637, referred to as nor-MDP), N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-sn-glycero-3-hydroxyphosphoryloxy)-ethylamine (CGP 19835A, referred to as MTP-PE), and RIBI, which contains three components extracted from bacteria, monophosphoryl lipid A, trehalose dimycolate and cell wall skeleton (MPL+TDM+CWS) in a 2% squalene/Tween 7 80 emulsion. Adjuvants may also include MF59 or sub-units of Cholera toxin which may be mutated or fragments thereof, or DNA itself which is an adjuvant known in the art. The effectiveness of an adjuvant may be determined by measuring the amount of antibodies directed against an immunogenic polypeptide containing an antigenic sequence resulting from administration of this polypeptide in vaccines which are also comprised of the various adjuvants.

In a further aspect of the invention there is provided an attenuated vaccine comprising a

normally pathogenic mycobacteria which harbours an attenuating mutation in the *mpa* gene or one or more of its homologues.

The mycobacteria may be used in the form of a killed bacteria or as a live attenuated vaccine. There are advantages to using a live attenuated vaccine. If the whole live organism is used rather than dead cells or selected cell components which may exhibit modified or denatured antigens, protein antigens in the outer membrane will tend to maintain their tertiary and quaternary structures and therefore the potential to illicit a good protective long term immunity should be higher.

The term attenuating mutation and the like refers to any genetic lesion in a gene which renders the gene non-functional. This may be, for example, by deletion of all or part of the gene, point mutation in the coding sequence resulting in a truncated gene product unable to carry out the normal function of the gene, or insertion or interruption of the gene by a nucleotide element which prevents the gene product being coded or causes any gene product generated to be such that it cannot carry out the normal function of the gene.

A mutation introduced into an attenuated bacterium of the invention would generally be a non-reverting attenuating mutation. Non-reverting means that for practical purposes the probability of the mutated gene being restored to its normal function is small, for example, less than 1 in 10^6 preferably less than 1 in 10^9 or more preferably less than 1 in 10^{12}. The attenuated mycobacteria of the invention may be isolated form. This is usually desirable when the bacteria are to be used for the purposes of vaccination. The term "isolated" means that the bacteria are in a form in which they can be cultured, processed or otherwise used in a form which can be readily identified and in which it is substantially uncontaminated by

other bacterial strains, for example, non-attenuated parent strains or unrelated bacterial strains. The term "isolated bacterium" thus enco*mpa*sses cultures of a bacterial mutant of the invention, for example, in the form of colonies on a solid medium or in the form of a liquid culture, as well as frozen or dried preparations of the strain.

In a preferred aspect the attenuated mycobacterium is formed by insertion of an insertion element into the *mpa* gene. The insertion element may consist of a sequence of one or more, preferably ten or more, nucleotides. In preferred attenuated mycobacteria the *mpa* gene is attenuated by insertion of a known insertion element, for example, an IS*21*-like element, more preferably by insertion of the insertion element IS*1612* or an insertion element at least 75% homologous to it. Most preferably the insertion of the IS*1612* element or its homologue occurs at the target site 5'-ATGCAC –3' between nucleotides 202 and 209 of the *mpa* sequence. The mycobacteria which is attenuated is preferably *Mptb*.

The attenuated mycobacterium may further comprise at least one additional mutation. This may be a mutation in a gene responsible for the production of products essential to bacterial growth which are absent in a human or animal host. For example, mutations to the gene for aspartate semi-aldehyde dehydrogenase (*asd*) have been proposed for the production of attenuated strains of *salmonella*. The *asd* gene is described further in Gene (1993) 129, 123-128. A lesion in the *asd* gene, encoding the enzyme aspartate *B*-semialdehyde dehydrogenase would render the organism auxotrophic for the essential nutrient diaminopelic acid (DAP), which can be provided exogenously during bulk culture of the vaccine strain. Since this compound is an essential constituent of the cell wall for gram-negative and some gram-positive organisms and is absent from mammalian or other vertebrate tissues, mutants would undergo lysis after about three rounds of division in such

tissues. Analogous mutations may be made to the attenuated mycobacteria of the invention. Such mutations may also include disabling genes of the GS element including *gsc* and *gsd*.

In addition or in the alternative, the attenuated mycobacteria may carry a *rec*A mutation. The *rec*A mutation knocks out homologous recombination - the process which is exploited for the construction of the mutations. Once the *rec*A mutation has been incorporated the strain will be unable to repair the constructed deletion mutations. Such a mutation will provide attenuated strains in which the possibility of homologous recombination to with DNA from wild-type strains has been minimized. *Rec*A genes have been widely studied in the art and their sequences are available. Further modifications may be made for additional safety.

The invention further provides a process for preparing a vaccine composition comprising an attenuated bacterium according to the invention process comprises (a) inoculating a culture vessel containing a nutrient medium suitable for growth of said bacterium; (b) culturing said bacterium; (c) recovering said bacteria and (d) mixing said bacteria with a pharmaceutically acceptable diluent or carrier.

The *mpa* gene can also be inactivated by methods known in the art including other transposon mutagenesis, and allelic exchange.

Attenuated mycobacterial strains according to the invention may be constructed using recombinant DNA methodology which is known per se. In general, bacterial genes may be mutated by a process of targeted homologous recombination in which a DNA construct containing a mutated form of the gene is introduced into a host bacterium which it is

desired to attenuate. The construct will recombine with the wild-type gene carried by the host and thus the mutated gene may be incorporated into the host genome to provide a bacterium of the present invention which may then be isolated.

The mutated gene may be obtained by introducing deletions into the gene, e.g by digesting with a restriction enzyme which cuts the coding sequence twice to excise a portion of the gene and then relegating under conditions in which the excised portion is not reintroduced into the cut gene. Alternatively frame shift mutations may be introduced by cutting with a restriction enzyme which leaves overhanging 5' and 3' termini, filling in and/or trimming back the overhangs, and religating. Similar mutations may be made by site directed mutagenesis. These are only examples of the types of techniques which will readily be at the disposal of those of skill in the art.

Various assays are available to detect successful recombination. In the case of attenuations which mutate a target gene necessary for the production of an essential metabolite or catabolite compound, selection may be carried out by screening for bacteria unable to grow in the absence of such a compound. Bacteria may also be screened with antibodies or nucleic acids of the invention to determine the absence of production of a mutated gene product of the invention or to confirm that the genetic lesion introduced - e.g. a deletion - has been incorporated into the genome of the attenuated strain. In addition, GPL sugars may be extracted and analysed by high pressure liquid chromatography and other methods such as TLC known in the art.

The concentration of the attenuated strain in the vaccine will be formulated to allow convenient unit dosage forms to be prepared. Concentrations of from about 10^4 to 10^9

bacteria per ml will generally be suitable, e.g. from about 10^5 to 10^8 such as about 10^6 per ml. Live attenuated organisms may be administered subcutaneously or intramuscularly at up to 10^8 organisms in one or more doses, e.g from around 10^5 to 10^8, e.g about 10^6 or 10^7 organisms in a single dose.

The vaccines of the invention are conventionally administered parenterally, by injection, for example, either subcutaneously or intramuscularly. Additional formulations which are suitable for other modes of administration such as ingestion, rectal installation, or mucosally by inhalation. For suppositories, traditional binders and carriers may include, for example, polyalkylene glycols or triglycerides; such suppositories may be formed from mixtures containing the active ingredient in the range of 0.5% to 10%, preferably 1% - 2%. Oral formulations include such normally employed excipients as, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate, and the like. These compositions take the form of solutions, suspensions, tablets, pills, capsules, sustained release formulations or powders and contain 10% - 95% of active ingredient, preferably 25% - 70%. Oral formulations may include enteric coatings so that vaccines are preferentially released in the small intestine or colon.

The proteins may be formulated into the vaccine as neutral or salt forms. Pharmaceutically acceptable salts include the acid addition salts (formed with free amino groups of the peptide) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids such as acetic, oxalic, tartaric, maleic, and the like. Salts formed with the free carboxyl groups may also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such

organic bases as isopropylamine, trimethylamine, 2-ethylamino ethanol, histidine, procaine, and the like.

The vaccines are administered in a manner co*mpa*tible with the dosage formulation, and in such amount as will be prophylactically and/or therapeutically effective. The quantity to be administered, which is generally in the range of 500μg to 5mgs of antigen per dose, depends on the subject to be treated, capacity of the subject's immune system to synthesize antibodies, mode of administration and the degree of protection desired. Precise amounts of active ingredient required to be administered may depend on the route of administration or on judgement of the practitioner and may be peculiar to each subject.

The vaccine may be given in a single dose schedule, or preferably in a multiple dose schedule. A multiple dose schedule is one in which a primary course of vaccination may be with 1-10 separate doses, followed by other doses given at subsequent time intervals required to maintain and or reenforce the immune response, for example, at 1-4 months for a second dose, and if needed, a subsequent dose(s) after several months. The dosage regimen will also, at least in part, be determined by the need of the individual and be dependent upon the judgement of the practitioner.

The vaccines of the invention may be administered to recipients to treat established diseases or in order to protect them against diseases caused by the corresponding wild type mycobacteria such as inflammatory diseases such as Crohn's disease or sarcoidoses in humans of Johne's disease in animals. The vaccine may be administered by any suitable route as described above.

The invention is now illustrated by the following Examples which should be construed as non-limiting.

Examples

Characterisation of the *mpa* gene.

PCR performed on *Mptb* DNA using a number of polynucleotide primers derived from the sequence of GS within *M.avium* sub sp. *silvaticum* (PCT/GB96/03221) failed to yield products with *Mptb.* Although the first 209 nucleotides of *mpa* are present in *M.avium* subsp.*silvaticum*, the existence of an open reading frame was not apparent because the sequence was truncated in *M.avium* subsp.*silvaticum* by the IS*21*-like element, now designated IS*1612*. Primers including those from within IS*1612* that were predicted to produce an amplification product and disclose further DNA sequences from within *Mptb,* if this were the same as *M.avium* subsp.*silvaticum*, did not work. The genomic organisation of *Mptb* in this region was found to be very different from that of *M.avium* subsp.*silvaticum.* Furthermore, the full extent of the IS*21*-like element (IS*1612*) in *M.avium* subsp.*silvaticum* remained unknown. When further experimentation disclosed the full sequence of IS1612 in *M.avium* subsp.*silvaticum*, primers selected on the basis of this new information again failed to give an amplification product, and reveal further DNA sequence of *Mptb.* It became clear that a DNA sequencing strategy for this region of *Mptb,* down stream of the GS element known in the art, based upon the DNA sequence information from *M.avium* subsp.*silvaticum* would not work.

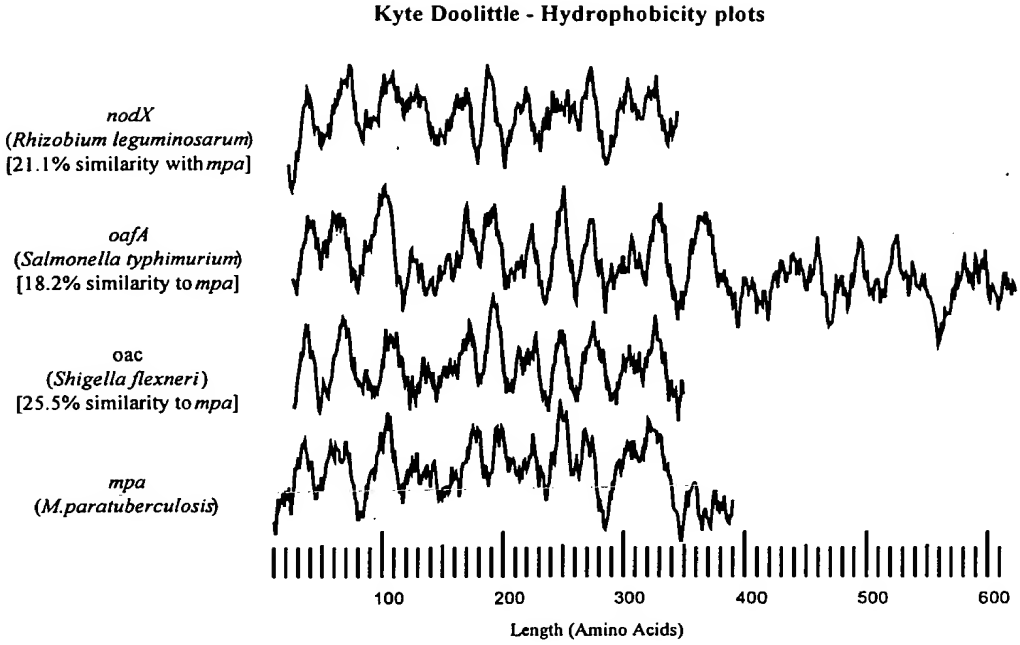
In a new strategy, we used the enzymes *Bam* HI, *Sal*I, *Pst*I and XhoI and a variety of

primers selected from the last known 150bp of DNA sequence in this region of *Mptb*, to obtain amplification products from *Mptb* DNA using inverse PCR. This involved recircularisation of genomic digest and subsequent amplification by a set of 'inward' and 'outward' primers. The results were at first difficult to assess because of the production of a number of non-specific amplification products. In several experiments using this strategy however, one band of 280 bp was consistently obtained. This suggested that at least one of the primers in the reaction was specific for further sequence of *Mptb*. This 280bp product was therefore selected from amongst the other amplification products for a further round of cloning and DNA sequencing. This resulted in our obtaining 250bp of sequence from within the *mpa* gene of *Mptb*. Further work revealed that *Mptb* did not contain an IS*21*-like element (IS*1612*) found in GS in *M.avium* subsp.*silvaticum*. It became clear that the *mpa* gene existed downstream of the GS element in *Mptb*. Using inverse PCR the entire sequence of the *mpa* gene (SEQ ID No. 1) and the *mpa* polypeptide which it encodes (SEQ ID No.2) were obtained and verified by sequencing both strands. Comparison of the sequence of the polypeptide in SEQ ID No. 2 with existing animo acid sequences available in the databases, demonstrated homologies which revealed the function of the *mpa* gene product to be an acetyl transferase closely linked with pathogenicity in other organisms.

Functional characterisation of *mpa*.

Sequence data for *mpa* was compared against known sequences using the BLASTN 1.4.11 [24-Nov-97] [Build 24-Nov-97] search facility at NCBI (website: ncbi.nlm.nih.gov/ genbank/) on GenBank Flat File release 105.0 (release date 15.02.98). Homologues to *mpa* were found in acetylases from *S.typhimurium, S.flexneri* and *R.leguminosarum*. The function of this ORF as an acetylase is strengthened by its very high similarity in

hydrophobicity plots of homologues.



In each of these cases the acetylase gene is crucial for determining the host specificity of the organism. Acetylation of lipopolysaccharide O-antigen by the *oac* gene, carried by a strain specific bacteriophage, converts *S.flexneri* to the more virulent 06 serotype (Clark CA, Beltrame J & Manning PA. Gene **107**: 43-52. 1991). Critical mutations or deletions in the *oafA* gene, which causes the acetylation of the exported O-antigen lipopolysaccharide in *S.typhimurium,* destroys an epitope crucial for O5 serotype antibody binding and decreases the virulence of this bacteria for host cells (Slauch JM, Lee AA, Mahan MJ & Mekalanos JJ. J.Bacteriol. **178**: 5904-5909.1996). An *mpa* homologue *nodX* is not present in all species of *Rhizobium,* however when present it causes the nod to be acetylated. This specific modification allows the strain to increase the range of hosts that it can stimulate to nodulate. Nod factors consist of polysaccharides that contain glucose, galactose and fucose residues which are methylated and sometimes acetylated. In mycobacteria also, acetylation of GPL's can be important for host recognition. De-acetylation of the terminal fucose of GPL's from

Mavium MAC serotype 9 abolishes rabbit anti-serotype 9 antibody agglutination indicating that acetylation plays a key role in strain variability amongst bacteria and may be directly attributable as a virulence factor. This is not the case for *Mycobacterium avium* subsp. *avium* serotype 2 GPL. Acetylation of rhamnose in GPL's of *Mycobacterium smegmatis* induces resistance to mycobacteriophage D4 by inducing conformational changes that destroy the phage attachment site. Thus in *Mptb*, the *mpa* gene is responsible for modification of terminal sugar residues in *Mptb* GPL's which are critically important in determining cell surface recognition and receptor binding and are important determinants of pathogenicity. *M.avium* subsp.*silvaticum* differs from *Mptb* in this important respect because the *mpa* gene in *M.avium* subsp.*silvaticum* is knocked-out by the insertion of IS*1612*. The use of IS*1612* using standard transfection methodologies to knock-out *mpa* in *Mptb* produces an attenuated mutated strain. The introduction of such specific mutations, including the *mpa* gene, is of use in the provision of attenuated vaccine strains of *Mptb* to increase the resistance of animals and humans to infections caused by this organism. Transfection of the *mpa* gene from *Mptb* into a recipient organism or eukaryotic cell particularly accompanied by the simultaneous transfection of GS will confer on the recipient the property of synthesising cell surface structures which mimic those of *mpa* and thus also serve as a vaccine to increase the resistance of animals and humans to *Mptb* infections.

Transformation of another bacterium by *mpa*.

Mpa specific primers are designed with concatenated SacI sites in p12 and a PstI site in p13 at the 3' ends (P12:AGCGAGCTCACGTGACTGAAGCC ; P13 :GCTCTGCAGCCGGAA CACAACGC). A 1371 bp PCR product is amplified, purified by gel electrophoresis and

cleaned with Qiagen gel clean column. The product is then cloned into T-Vector (Promega) and transformed into E.coli. The plasmid is grown up, purified and a Pst1 insert fragment removed and purified as before. This is subcloned into a mycobacterial shuttle vector. This vector contains a mycobacterial ori, *E.coli* ori, hygromycin antibiotic marker and the hsp60 promoter immediately upstream of the insertion locus. The resulting construct is transformed into *E.coli* and purified to 1mg/ml. This vector is then sequenced to check that the PCR step and cloning steps do not introduce errors into the sequence and that the sequence is inserted in the correct orientation in the vector. *M.smegmatis* (strain MC^2155) is grown into exponential phase, washed twice in 10% glycerol and diluted to OD260:0.5 in TE x1 (Tris 10mM EDTA 0.1mM). 1µg of plasmid is added and the mixture pulsed at 1000Ω:2.5kV:25fD using a BioRad electroporation unit. Cells are recovered in 500µl SOC (2g Tryptone, 0.5g Yeast extract, 1ml 1M NaCl, 0.25ml 1MKCl, 1ml Mg Salts, 1ml 2M Glucose) at 37°C for 3 hours and then plated onto Middlebrooks 7H11 : 45µg/ml Hygromycin selection plates. Transformants are selected and *mpa* presence checked by *mpa* specific P12/P13 PCR. Expression of *mpa* is also checked by making total mRNA preps of transformants in exponential growth phase. This is done by pelleting a 4 day 30ml culture (in Middlebrooks 7H11 : 45µg/ml Hygromycin broth) at 3,000 xg for 20 minutes. This is resuspended in 200µl RNAse free water and transferred to a ribolyser tube containing silica beads. 500µl of DSA solution (Divolab No.1:9.6ml, 500mM Na Acetate(pH4.0) 24ml, RNAse free water 66.4ml) 500µl Acid Phenol (Water saturated phenol with Na Acetate at pH4.0) and 100µl chloroform/isoamyl alcohol (24:1). This is then ribolysed at a speed setting of 6.5 for 45secs. The tubes are then microfuged for 10 minutes at 13,000 xg. Total mRNA is extracted with equal volume of chloroform/isoamyl alcohol and then precipitated at -70°C with isopropanol for 2 hours. This is then microfuged at 13,000 xg for 20 minutes dried and resuspended in RNAase free water. Samples are treated with DNAase and then

cDNA produced using the P12 primer as template with SuperscriptII Reverse Transcriptase (GibcoBRL). PCR using P12/P13 primers is then performed showing bands of the correct size. Controls without DNAase and with RNAase H treatment are performed in parallel. This demonstrates transcribed copies of *mpa* present in *M.smegmatis*. Translation is revealed by taking exponential growth cultures of transformants and whole cell lysates using 1 minute sonication in 2% SDS-PAGE buffer. These whole cell protein extracts are then electrophoresed on 1% polyacrylamide gels and western blotted onto nylon membranes. These are then hybridised with rabbit raised, *mpa* peptide antisera and developed with anti-rabbit HRP conjugate/ECL peroxide system.

Deactivation of *mpa* in *Mptb* using IS*1612*

General transposon mutagenesis of *Mptb* can be achieved using Tn-Mut.Vectors. These vectors contain the insertion sequence, a kanamycin selection marker, shuttle vector origins of replication (including a thermosensitive Myc-ori) and the counterselectable suicide marker *sac*B, and are of proven ability in the mutagenesis of BCG, *M.tuberculosis*, *M.phlei* and *M.smegmatis*. Selected drug resistant transformants may be screened for interruption of the *mpa* gene. However, we have shown that *mpa* in *M.avium* subsp.*silvaticum* is interrupted and inactivated by the insertion at a specific site of IS*1612* and that *Mptb* does not contain IS*1612* and its *mpa* gene is intact. IS*1612* cloned from a clinical isolate of *M.avium* subsp.*silvaticum* can be used in a suitable construct for the specific inactivation of *mpa* in *Mptb*. Previous studies have shown that IS*21*-like elements resembling IS*1612* can have a selected DNA marker inserted into a region immediately downstream of the *ist*B gene without losing transpositional activity. Using primer mutagesesis a unique restriction site is created 4bp downstream of IS*1612*/ *ist*B. A hygromycin selection marker is inserted

into this site and the resultant transposon substituted into the Tn-Mut vector in place of IS*1096*. This vector is introduced by electroporation into *Mptb* or a pathogenic *M.avium* containing *mpa*, and successful transformants isolated using the selection and counterselection markers and thermotolerant conditions 30°C. Cultures are subbed at intervals into 2% sucrose media and grown at a non-permissive temperature (37°C). The selection marker is co-transferred with the insertion element and selects for transpositional events whilst the temperature suppresses plasmid replication, inducing plasmid leaching and killing plasmid positives via *sac*B expression. DNA extracted from clones will be screened using a gene specific primer and an insertion sequence specific primer. Cloned PCR products will be sequenced and checked against known target sequence data. Clones will then be cultured and RFLP analysis performed using insertion sequence specific probes to determine insertion frequency. Clones with single insertions in the *mpa* gene are selected for use as candidate vaccine strains with attenuated pathogenicity.

SEQUENCE LISTINGS

Sequence ID No.1

Total bases = 1,335

```
1     GTGACTGAAGCCAATGAGTGCAACTCGGCGTCGCGAAAGGTTTCAGTCGC
51    GGTTGAGCAAGACACCGCAAGACTACTGGAGTGCGTGCACAAGCGCCCCC
101   AGCTCGCGGCTGAAAGCGGATGCAAAGGGGTTCGAAGCTTGAGCAACATG
151   CGAAGGGGAGAACGGCCTATGAGCCTGGGACAGGTTTTCGACCCGCGCGC
201   GAATGCACTTAATGCGTGGCGCTTGGTGTTGGCGAGCGGGGTGATCCTAT
251   GGCATTCGTTTCCGCTCACTGGACGTATGCCGTGGGCGCCGTTCGTCCAG
301   TTGCTTGGCCTTGGATGCGTTGATGGTTTCTTTGCGGTCTCCGGCTATCT
351   CATCGTCTCGAGCTGGCTTCGCAACCCGCATCCCGCCCAATACTTCACCG
401   CTCGATGTCTTCGTATTCTCCCGGGTCTGTGGATCTGTCTCATCTTGACG
451   GCGTTTGTCATCGCTCCGATAGGTGTGGGCGCCCAGGGCGGTTCGGCCGC
501   GAAACTACTGATGTCCGGCGCTCCGATCGAGTATGTGCTAAAAGACAGTG
551   CGGTTTGGATGGTTAAGTTCGATATCGGTGGCACACCTCGCGATATTCCA
601   GTTGCGGGTATTTGGAACGGTTCTCTGTGGACATTGGGTTGGGAGGTGCT
651   TTGCTATATCGGCGTAGCAGTATTTGGTATGCTCGGACTTCTTAGTCGCC
701   GTTGGTTCGTTCCAGGGATATTGATCCTGGCGCTGTCGTGGTCGGTGTTC
751   TTGCCGGCATGGGGCGGAATACACGCGATCGCCTCCAATGCTGCGCGATT
801   CGCTGTGATGTTTTCGGCCGGAGCGTTGCTGTATCAATTCCGTAACGTGA
851   TTCCGGCTCGGTGGTCCTTCGTTGCCGTCGGCCTCATTATCGTTGTGGTT
901   TCCTCTGCCGTGCTGCCGGACTACCGGTTGGTGGCGGCCCTTCCGATGGC
951   GTACCTAATCATCGCTTCGGGTTCGCTCATCCACAATCAAAGGATGAGGT
1001  TCCGCACCGATCTATCCTATGGAGTATATATTTATGCGTTTCCAATTCAG
1051  CAAGTGCTGGTCCTGTGTGGATTCGCCGAGATAAATCCAATCGCTTTCTG
1101  CGCGATTTCTGTCGCAGCTATTTTGCCGCTCGCCGCGCTCAGTTGGTTCT
1151  TGGTCGAGAAACCTGCGTTGTCCTGGAAGAGTCGTCTCCGGCGGAAAAAC
1201  AGTTCAATTGCGCTAGCCAATATGGAAGATGGTGGATCAGTCGGCCGCTC
1251  AAATGACATTCCCGGAAGGCGGGCCCGCTTTATTGGCGAGAAAGCCGAAG
1301  ATCCTCCCGCGCCGAGCCCAAGACCGGCTTTGTAA
```

Sequence ID No.2

Total Residues = 444

```
1     VTEANECNSASRKVSVAVEQDTARLLECVHKRPQLAAESGCKGVRSLSNM
51    RRGERPMSLGQVFDPRANALNAWRLVLASGVILWHSFPLTGRMPWAPFVQ
101   LLGLGCVDGFFAVSGYLIVSSWLRNPHPAQYFTARCLRILPGLWICLILT
151   AFVIAPIGVGAQGGSAAKLLMSGAPIEYVLKDSAVWMVKFDIGGTPRDIP
201   VAGIWNGSLWTLGWEVLCYIGVAVFGMLGLLSRRWFVPGILILALSWSVF
251   LPAWGGIHAIASNAARFAVMFSAGALLYQFRNVIPARWSFVAVGLIIVVV
301   SSAVLPDYRLVAALPMAYLIIASGSLIHNQRMRFRTDLSYGVYIYAFPIQ
351   QVLVLCGFAEINPIAFCAISVAAILPLAALSWFLVEKPALSWKSRLRRKN
401   SSIALANMEDGGSVGRSNDIPGRRARFIGEKAEDPPAPSPRPAL
```

Sequence ID No. 3 = IS1612 positive strand.

```
1     ATGCACTGTCAATGGCCAAGTAGAAGTCCCCGCTGGTGGCCAGCAGAAGT
51    CCCCACTCCGCTGCGGGTGGTTGGCTAATTCTTGGCGGCTCCCTTCTTGT
101   GGTCGGCGTGGCGCATCCGGTAGGACTCGCCGGAGGTGACGACGATGCTG
151   GCGTGGTGCAGCAGCCGATCGAGGATGCTGGCGGCGGTGGTGTGCTCGGG
201   CAGGAATCGCCCCCATTGTTCGAAGGGCCAATGCGAGGCGATGGCCAGGG
251   AGCGGCGCTCGTAGCCGGCAGCCACGAGCCGGAACAACAGTTGAGTCCCG
301   GTGTCGTCGAGCGGGGCGAAGCCGATCTCGTCCAAGATGACCAGATCCGC
351   GCGGAGCAGGGTGTCGATGATCTTGCCGACGGTGTTGTCGGCCAGGCCGC
401   GGTAGAGGACCTCGATCAGGTCGGCGGCGGTGAAGTAGCGGACTTTGAAT
451   CCGGCGTGGACGGCAGCGTGCCCGCAGCCGATGAGCAGGTGACTTTTGCC
501   CGTACCAGGTGGGCCAATGACCGCCAGGTTCTGTTGTGCCCGAATCCATT
551   CCAGGCTCGACAGGTAGTCGAACGTGGCTGCGGTGATCGACGATCCGGTG
601   ACGTCGAACCCGTCGAGGGTCTTGGTGACCGGGAAGGCTGCGGCCTTGAG
651   ACGGTTGGCGGTGTTGGAGGCATCGCGGGCAGCGATCTCGGCCTCAACCA
701   ACGTCCGCAGGATCTCCTCCGGTGTCCAGCGTTGCGTCTTGGCGACTTGC
751   AACACCTCGGCGGCGTTGCGGCGCACCGTGGCCAGCTTCAACCGCCGCAG
801   CGCCGCGTCAAGGTCAGCAGCCAGCGGTGCCGCCGAGGACGGTGCCACCG
851   GCTTGGCAGCGGTGGTCATGAGGCCGTCCCGTCGGTGGTGTTGATCTTGT
901   AGGCCTCCAACGAGCGGGTCTCGACGGTGGGCAGATCGAGCACGAGTGCG
951   TCGCCGGCGGGGCGGGGTTGTGGGGTGCCGGCGCCGGCGGCCAGGATCGA
1001  GCGCACGTCGGCAGCGCGGAACCGGCGAAACGCAACCGCCCGGCGCAGCG
1051  CGTCAATCAAAGCCTGTTCGCCGTGGGCGGCGCCAAGGCCGAGCAGAATG
1101  TCGAGTTCGGATTTCAGTCGGGTGTTGCCGATCGCAGCAGCACCGACGAG
1151  GAACTGCTGCGCTTCGGTTCCCAATGCGCAGAATCGTTTCTCTGCTTGGG
1201  TTTTCGGGCGAGGACCACGCGAGGGTGCGGGTCTGGGTCCGTCGTAGTGT
1251  TCATCGAGGATGGACACCTCACCTGGGCTGACGAGCTCGTGCTCGGCCAC
1301  GATCACACCGGTCGCAGGTTCCAACAGGATCAGGGCGCCATGATCGACCA
1351  CCACCGCCACGGTGGCACCGACGAGCCGCTGAGGCACCGAGTAACGAGCT
1401  GAGCCGTAACGGATGCACGAGAGGCCGTCGACCTTACGGCGCACCGACCC
1451  CGAGCCGATCGTCGGCCGCAGCGAGGGCAGCTCCCTCAAGACGGTGCGCT
1501  CGTCAACCAAGCGATCGTTGGGCACGGCGCAGATCTCCGAGTGGACCGTG
1551  GCATTGACCTCGGCGCACCATAGTTGCGCCTGGGCGTTGAGGGCACGTAG
1601  GTCGACCTGCTCACCGGCTAACGCAGCTTCGGTCAGCAGCGGCACCGCAA
1651  GGTCGTCCTGAGCGTAGCCACAGAGGTTCTCCACGATGCCCTTCGATTGC
1701  GGATCCGCACCGTGGCAGAAGTCCGGAACGAAGCCATAGTGGGACGCGAA
1751  TCGCACATAATCCGGTGTTGGAACAACAACATTGGCGACGACACCACCTT
1801  TGAGGCAGCCCATCCGGTCGGCCAGGATCTTGGCCGGAACCCCACCGATC
1851  GCCTCGAGGGCTTCGGCTATCATCGCCTGCGTGGTCGAGGCTTTCTCGTC
1901  GGCGGCGAACCGCTCAAACCGCCACCGCGAATAGGCCAGCTCCGCGCATA
1951  ACACCATCAGCCCCGGTGCCGCTTCGGCCCAATCCATCACCAGATAGTCA
2001  CCGGGTGACCAGACCGCCGGACGGCGTTGATGCCGGTTAGCGTTGCGCCA
2051  CCATACTTCCTGCTCGGCTACCAGGCGGCGGAAGTTACGGGCCGAGCCCT
2101  GATACCCGGCAGCTCGGGCGATCGGCAGCATCCGCTTCGCCGACATCTTG
2151  CCGTGTGATTTCTCGACTCGGGTGGCGACTAGATCGGTGAACGCGTCGAG
2201  GTTGCGTGGCCGTGGTTCCCGCGGGGGCGCGCCACCGGCCTCGGCCCGCT
2251  CGATGACCCGCTTGACCGTCTTGTGCGTACTACCGCACAGCTCGGCCGCG
2301  CCGCGATACGACCCGACCTGGTGATACGCCGAAATGATGTTCATACGCTC
2351  CCTTGCAGACTTCAATAGAGCTCCCTGGGCGGTGATCAAGTGACAGTTGG
2401  CGCTATCACCGTCACCGCCCAGGCCCTCAGCTCCCGGAAAAGACACGACG
2451  AGCCCGCTAAGGAGTGGGGACTTCTACCTGGCCACCAGTGGGGACTTCCT
2501  ACTGGCCACAGATGGGGACTTTCTCATGGCCATGGACATGCAC
```

(Total = 2543bp)

Sequence ID No.4 = IS1612 negative strand to SEQ ID No.3.

```
1     GTGCATGTCCATGGCCATGAGAAAGTCCCCATCTGTGGCCAGTAGGAAGT
51    CCCCACTGGTGGCCAGGTAGAAGTCCCCACTCCTTAGCGGGCTCGTCGTG
101   TCTTTTCCGGGAGCTGAGGGCCTGGGCGGTGACGGTGATAGCGCCAACTG
151   TCACTTGATCACCGCCCAGGGAGCTCTATTGAAGTCTGCAAGGGAGCGTA
201   TGAACATCATTTCGGCGTATCACCAGGTCGGGTCGTATCGCGGCGCGGCC
251   GAGCTGTGCGGTAGTACGCACAAGACGGTCAAGCGGGTCATCGAGCGGGC
301   CGAGGCCGGTGGCGCGCCCCCGCGGGAACCACGGCCACGCAACCTCGACG
351   CGTTCACCGATCTAGTCGCCACCCGAGTCGAGAAATCACACGGCAAGATG
401   TCGGCGAAGCGGATGCTGCCGATCGCCCGAGCTGCCGGGTATCAGGGCTC
451   GGCCCGTAACTTCCGCCGCCTGGTAGCCGAGCAGGAAGTATGGTGGCGCA
501   ACGCTAACCGGCATCAACGCCGTCCGGCGGTCTGGTCACCCGGTGACTAT
551   CTGGTGATGGATTGGGCCGAAGCGGCACCGGGGCTGATGGTGTTATGCGC
601   GGAGCTGGCCTATTCGCGGTGGCGGTTTGAGCGGTTCGCCGCCGACGAGA
651   AAGCCTCGACCACGCAGGCGATGATAGCCGAAGCCCTCGAGGCGATCGGT
701   GGGGTTCCGGCCAAGATCCTGGCCGACCGGATGGGCTGCCTCAAAGGTGG
751   TGTCGTCGCCAATGTTGTTGTTCCAACACCGGATTATGTGCGATTCGCGT
801   CCCACTATGGCTTCGTTCCGGACTTCTGCCACGGTGCGGATCCGCAATCG
851   AAGGGCATCGTGGAGAACCTCTGTGGCTACGCTCAGGACGACCTTGCGGT
901   GCCGCTGCTGACCGAAGCTGCGTTAGCCGGTGAGCAGGTCGACCTACGTG
951   CCCTCAACGCCCAGGCGCAACTATGGTGCGCCGAGGTCAATGCCACGGTC
1001  CACTCGGAGATCTGCGCCGTGCCCAACGATCGCTTGGTTGACGAGCGCAC
1051  CGTCTTGAGGGAGCTGCCCTCGCTGCGGCCGACGATCGGCTCGGGGTCGG
1101  TGCGCCGTAAGGTCGACGGCCTCTCGTGCATCCGTTACGGCTCAGCTCGT
1151  TACTCGGTGCCTCAGCGGCTCGTCGGTGCCACCGTGGCGGTGGTGGTCGA
1201  TCATGGCGCCCTGATCCTGTTGGAACCTGCGACCGGTGTGATCGTGGCCG
1251  AGCACGAGCTCGTCAGCCCAGGTGAGGTGTCCATCCTCGATGAACACTAC
1301  GACGGACCCAGACCCGCACCCTCGCGTGGTCCTCGCCCGAAAACCCAAGC
1351  AGAGAAACGATTCTGCGCATTGGGAACCGAAGCGCAGCAGTTCCTCGTCG
1401  GTGCTGCTGCGATCGGCAACACCCGACTGAAATCCGAACTCGACATTCTG
1451  CTCGGCCTTGGCGCCGCCCACGGCGAACAGGCTTTGATTGACGCGCTGCG
1501  CCGGGCGGTTGCGTTTCGCCGGTTCCGCGCTGCCGACGTGCGCTCGATCC
1551  TGGCCGCCGGCGCCGGCACCCCACAACCCCGCCCCGCCGGCGACGCACTC
1601  GTGCTCGATCTGCCCACCGTCGAGACCCGCTCGTTGGAGGCCTACAAGAT
1651  CAACACCACCGACGGGACGGCCTCATGACCACCGCTGCCAAGCCGGTGGC
1701  ACCGTCCTCGGCGGCACCGCTGGCTGCTGACCTTGACGCGGCGCTGCGGC
1751  GGTTGAAGCTGGCCACGGTGCGCCGCAACGCCGCCGAGGTGTTGCAAGTC
1801  GCCAAGACGCAACGCTGGACACCGGAGGAGATCCTGCGGACGTTGGTTGA
1851  GGCCGAGATCGCTGCCCGCGATGCCTCCAACACCGCCAACCGTCTCAAGG
1901  CCGCAGCCTTCCCGGTCACCAAGACCCTCGACGGGTTCGACGTCACCGGA
1951  TCGTCGATCACCGCAGCCACGTTCGACTACCTGTCGAGCCTGGAATGGAT
2001  TCGGGCACAACAGAACCTGGCGGTCATTGGCCCACCTGGTACGGGCAAAA
2051  GTCACCTGCTCATCGGCTGCGGGCACGCTGCCGTCCACGCCGGATTCAAA
2101  GTCCGCTACTTCACCGCCGCCGACCTGATCGAGGTCCTCTACCGCGGCCT
2151  GGCCGACAACACCGTCGGCAAGATCATCGACACCCTGCTCCGCGCGGATC
2201  TGGTCATCTTGGACGAGATCGGCTTCGCCCCGCTCGACGACACCGGGACT
2251  CAACTGTTGTTCCGGCTCGTGGCTGCCGGCTACGAGCGCCGCTCCCTGGC
2301  CATCGCCTCGCATTGGCCCTTCGAACAATGGGGGCGATTCCTGCCCGAGC
2351  ACACCACCGCCGCCAGCATCCTCGATCGGCTGCTGCACCACGCCAGCATC
2401  GTCGTCACCTCCGGCGAGTCCTACCGGATGCGCCACGCCGACCACAAGAA
2451  GGGAGCCGCCAAGAATTAGCCAACCACCCGCAGCGGAGTGGGGACTTCTG
2501  CTGGCCACCAGCGGGGACTTCTACTTGGCCATTGACAGTGCAT
```

(Total = 2543bp)

Sequence ID No.5 = Amino acid sequence of polypeptide designated istA encoded by Seq ID No.4.

```
1    VSFPGAEGLGGDGDSANCHLITAQGALLKSARERMNIISAYHQVGSYRGA
51   AELCGSTHKTVKRVIERAEAGGAPPREPRPRNLDAFTDLVATRVEKSHGK
101  MSAKRMLPIARAAGYQGSARNFRRLVAEQEVWWRNANRHQRRPAVWSPGD
151  YLVMDWAEAAPGLMVLCAELAYSRWRFERFAADEKASTTQAMIAEALEAI
201  GGVPAKILADRMGCLKGGVVANVVVPTPDYVRFASHYGFVPDFCHGADPQ
251  SKGIVENLCGYAQDDLAVPLLTEAALAGEQVDLRALNAQAQLWCAEVNAT
301  VHSEICAVPNDRLVDERTVLRELPSLRPTIGSGSVRRKVDGLSCIRYGSA
351  RYSVPQRLVGATVAVVVDHGALILLEPATGVIVAEHELVSPGEVSILDEH
401  YDGPRPAPSRGPRPKTQAEKRFCALGTEAQQFLVGAAAIGNTRLKSELDI
451  LLGLGAAHGEQALIDALRRAVAFRRFRAADVRSILAAGAGTPQPRPAGDA
501  LVLDLPTVETRSLEAYKINTTDGTAS
```

Coded by nucleotides 98 - 1678 of seq ID No 4.

(Total = 526 aa)

CLAIMS

1. A polynucleotide encoding a protein that comprises mycobacterium paratuberculosis acylase (*mpa*), or a fragment or homologue of said protein, said fragment or homologue having *mpa* activity.

2. A polynucleotide selected from:

(a) a polynucleotide comprising the entire nucleotide sequence set out in SEQ ID No. 1 or the complement thereof;

(b) a polynucleotide comprising a nucleotide sequence capable of hybridising to a fragment of the nucleotide sequence set out in SEQ ID No. 1, the fragment having the nucleotide sequence of nucleotides 210-1335 of SEQ ID No. 1;

(c) a polynucleotide comprising a nucleotide sequence capable of hybridising to the complement of a fragment of the nucleotide sequence set out in SEQ ID No. 1, the fragment having the nucleotide sequence of nucleotides 210-1335 of SEQ ID No. 1;

(d) a polynucleotide comprising a polynucleotide sequence which is degenerate as a result of the genetic code to the polynucleotide of SEQ ID No. 1 or a polynucleotide of (c); and

(e) a polynucleotide having at least 80% homology to the nucleotide sequence of SEQ ID No. 1.

3. A polynucleotide according to claim 2 which encodes a polypeptide having *mpa* activity.

4. A polynucleotide according to claim 1 or 3 wherein the *mpa* activity is the acetylation of cell wall components.

5. A polynucleotide probe or primer which comprises a fragment of at least 15 nucleotides of a polynucleotide selected from:

(b) a polynucleotide comprising a nucleotide sequence capable of hybridising to a fragment of the nucleotide sequence set out in SEQ ID No. 1, the fragment having the nucleotide sequence of nucleotides 210-1335 of SEQ ID No. 1;

(c) a polynucleotide comprising a nucleotide sequence capable of hybridising to the complement of a fragment of the nucleotide sequence set out in SEQ ID No. 1, the

fragment having the nucleotide sequence of nucleotides 210-1335 of SEQ ID No. 1; and

(d') a polynucleotide comprising a polynucleotide sequence which is degenerate as a result of the genetic code to a polynucleotide sequence of (c).

6. A polypeptide in substantially isolated form which is encoded by a polynucleotide of any one of claims 1 to 4.

7. A polypeptide in substantially isolated form which comprises the sequence set out in SEQ ID No. 2, or a polypeptide substantially homologous thereto which has *mpa* activity, or a fragment of the polypeptide of SEQ ID No. 2 which has *mpa* activity.

8. A polypeptide according to claim 7 which has the sequence set out in SEQ ID No. 2.

9. A polypeptide comprising at least 8 amino acids which is an immunogenic fragment of a polypeptide defined in claim 7 or 8 and which comprises an epitope which is specific to the pathogenicity of mycobacterial cells.

10. A vector comprising a polynucleotide as defined in any one of claims 1 to 4.

11. An expression vector comprising a polynucleotide as defined in any one of claims 1 to 4, operably linked to regulatory sequences capable of directing expression of said polynucleotide in a host cell.

12. An antibody capable of recognising a polypeptide as defined in any one of claims 6 to 9.

13. An antibody according to claim 12 which is a monoclonal antibody or a fragment thereof.

14. A method for detecting the presence or absence of a polynucleotide as defined in any one of claims 1 to 4 in a biological sample which method comprises:

(a) bringing a biological sample containing DNA or RNA into contact with a probe according to claim 5 under hybridising conditions; and

(b) detecting any duplex formed between the probe and nucleic acid in the sample.

15. A method of detecting the presence or absence of a polypeptide as defined in any one of claims 6 to 9 in a biological sample which method comprises:

(a) incubating the biological sample with an antibody according to claim 12 or 13 under conditions which allow for the formation of an antibody-antigen complex; and

(b) determining whether antibody-antigen complex comprising said antibody is formed.

16. A method of detecting the presence or absence of antibodies in a biological sample which method comprises:

(a) incubating a biological sample with a polypeptide according to any one of claims 6 to 9 comprising an epitope under conditions which allow for the formation of an antibody-antigen complex; and

(b) determining whether an antibody-antigen complex comprising said polypeptide is formed.

17. A method of detecting the presence or absence of cell mediated immune reactivity in an animal or human, to a polypeptide according to any one of claims 6 to 9 which method comprises:

(a) incubating a cell sample with a polypeptide according to any one of claims 6 to 9 comprising an epitope under conditions which allow for a cellular immune response; and

(b) detecting the presence of said cellular immune response in the incubate.

18. A test kit for detecting the presence or absence of a pathogenic mycobacterium in a sample which comprises a polynucleotide according to any one of claims 1 to 4 or a polypeptide according to any one of claims 6 to 9 or an antibody according to claim 12 or 13.

19. A pharmaceutical composition comprising (i) a polypeptide according to any one of claims 6 to 9 or a polynucleotide according to any one of claims 1 to 4 or an antibody according to claim 12 or 13 and (ii) a suitable carrier or diluent.

20. A polypeptide according to any one of claims 6 to 9 or a polynucleotide according to any one of claims 1 to 4 or an antibody according to claim 12 or 13, for use in the treatment, prevention or diagnosis of a disease caused by a mycobacterium.

21. A method of treating or preventing a mycobacterial disease in an animal or human caused by mycobacteria which express a polypeptide according to any one of claims 6 to 9, which method comprises administering to the animal or human an effective amount of said polypeptide.

22. A method of treating or preventing a mycobacterial disease in animals or humans caused by mycobacteria containing the nucleotide sequence of SEQ ID No. 1, which method comprises administering to the animal or human an effective amount of a polynucleotide according to any one of claims 1 to 4 or a vector according to claim 10 or 11.

23. A method according to claims 21 or 22 wherein the mycobacterial disease is Johne's disease or Crohn's disease.

24. A method according to any one of claims 21 to 23 for increasing the in vivo susceptibility of mycobacteria to antimicrobial drugs.

25. A vaccine composition comprising (i) a polypeptide according to any one of claims 6 to 9 or a polynucleotide according to any one of claims 1 to 4 or a vector according to claims 10 or 11 together with (ii) a pharmaceutically acceptable carrier or diluent.

26. A plasmid containing a polynucleotide sequence according to any one of claims 1 to 4 under the control of a promoter.

27. A plasmid according to claim 26 wherein the promoter is a CMV, MMLV, RSV or SV40 promoter.

28. A nucleic acid vaccine comprising (i) a plasmid as defined in claim 26 or 27 and (ii) a pharmaceutically acceptable carrier or diluent.

29. A vaccine according to claim 28 which further comprises a transfection agent.

30. A vaccine comprising (i) a polypeptide as defined in any one of claims 6 to 9, optionally linked to a hapten molecule, and (ii) a pharmaceutically acceptable carrier or diluent.

31. A non-pathogenic microorganism or a cell from a human or animal species prone to infection by *mpa*-containing mycobacteria comprising a component on its surface which has been modified by a polypeptide according to any one of claims 6 to 9.

32. A non-pathogenic microorganism or a cell from a human or animal species prone to infection by *mpa*-containing mycobacteria which has been transformed or transfected with a nucleic acid construct comprising a polynucleotide as defined in any one of claims 1 to 4 and 26 to 29.

33. A non-pathogenic microorganism or a cell from a human or animal species prone to infection by *mpa*-containing mycobacteria wherein the nucleic acid construct according to claim 32, further comprises a polynucleotide which encodes the polypeptides of the GS region of *Mptb*.

34. A non-pathogenic microorganism or a cell from a human or animal species prone to infection by *mpa*-containing mycobacteria according to claim 32 or 33 wherein the gene or genes present in the nucleic acid construct are expressed.

35. A vaccine comprising (i) a non-pathogenic microorganism or a cell from a human or animal species prone to infection by *mpa*-containing mycobacteria according to any one of claims 31 to 34 and (ii) a pharmaceutically acceptable carrier or diluent.

36. A non-pathogenic microorganism or a cell from a human or animal species prone to infection by *mpa*-containing mycobacteria comprising on its surface an antigenic determinant capable of being produced by the action of a polypeptide as defined in claim 7 or 8 and which is capable of eliciting antibodies which bind the surface of Mptb.

37. A normally pathogenic mycobacterium or pathogenic isolate thereof, whose pathogenicity is mediated in all or in part by the presence or expression of a polypeptide as defined in any one of claims 6 to 9, which mycobacterium or isolate harbours an attenuating mutation in the polynucleotide sequence as defined in any one of claims 1 to 4.

38. A vaccine comprising (i) a non-pathogenic microorganism or cell from a human or animal species as defined in claim 36 or a mycobacterium or isolate as defined in claim 37, and

(ii) a pharmaceutically acceptable carrier or diluent.

39. A vaccine according to claim 38 which comprises a mycobacterium or isolate wherein the attenuating mutation in the mycobacterium or isolate is mediated by the insertion of one or more nucleotides.

40. A polynucleotide insertion element selected from:

(a) a polynucleotide comprising the nucleotide sequence set out in SEQ ID Nos. 3 or 4;

(b) a polynucleotide comprising a nucleotide sequence capable of hybridising to a fragment of the nucleotide sequence set out in SEQ ID No. 3, the fragment having the nucleotide sequence of nucleotides 1856-2543 of SEQ ID No. 3;

(c) a polynucleotide comprising a nucleotide sequence capable of hybridising to a fragment of the nucleotide sequence set out in SEQ ID No. 4, the fragment having the nucleotide sequence of nucleotides 1-688 of SEQ ID No. 4;

(d) a polynucleotide comprising a polynucleotide sequence which is degenerate as a result of the genetic code to the polynucleotide of SEQ ID No. 4 or a polynucleotide of (b);

(e) a polynucleotide having at least 75% homology to the nucleotide sequence of SEQ ID No. 3; and

(f) a polynucleotide having at least 75% homology to the nucleotide sequence of SEQ ID No. 4.

41. A vaccine according to claim 39 wherein the sequence of one or more nucleotides is a sequence as defined in claim 40.

42. A polynucleotide probe or primer which comprises a fragment of at least 15 nucleotides of a polynucleotide selected from:

(b) a polynucleotide comprising a nucleotide sequence capable of hybridising to a fragment of the nucleotide sequence set out in SEQ ID No. 3, the fragment having the nucleotide sequence of nucleotides 1856-2543 of SEQ ID No. 3;

(c) a polynucleotide comprising a nucleotide sequence capable of hybridising to a fragment of the nucleotide sequence set out in SEQ ID No. 4, the fragment having

the nucleotide sequence of nucleotides 1-688 of SEQ ID No. 4; and

(d') a polynucleotide comprising a polynucleotide sequence which is degenerate as a result of the genetic code to a polynucleotide of (b).

43. A polypeptide in substantially isolated form which is encoded by a polynucleotide of claim 40.

44. A polypeptide comprising at least 8 amino acids which is an immunogenic fragment of the polypeptide defined in claim 43 and which comprises an istA epitope.

45. A vector comprising a polynucleotide as defined in claim 40.

46. An expression vector comprising a polynucleotide as defined in claim 40, operably liked to regulatory sequences capable of directing expression of said polynucleotide in a host cell.

47. A method for preparing a mycobacterium or pathogenic isolate as defined in claim 37 which method comprises transfecting animal or human isolate of an *mpa* containing pathogenic bacterium with a polynucleotide construct comprising a polynucleotide as defined in claim 40.

48. The method of claim 47 wherein transfection is effected by electroporation.

49. The method of claim 47 or 48 wherein the polynucleotide sequence as defined in claim 40 has the nucleotide sequence set out in SEQ ID No. 3 or 4.